








EX LIBRIS  
UNIVERSITATIS  
ALBERTENSIS

---

The Bruce Peel  
Special Collections  
Library





Digitized by the Internet Archive  
in 2025 with funding from  
University of Alberta Library

<https://archive.org/details/0162015205220>









University of Alberta

Library Release Form

Name of Author: Heather Lawrie Wilson

Title of Thesis: *SHANK3*, a Candidate Gene for the 22q13 Deletion Syndrome

Degree: Masters of Science

Year this Degree Granted: 2001

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.





University of Alberta

***SHANK3*, a Candidate Gene for the 22q13 Deletion Syndrome**

BY

**Heather Lawrie Wilson**



A thesis submitted to the Faculty of Graduate Studies and Research in  
partial fulfillment of the requirements for the degree of Master of Science.

In

Molecular Biology and Genetics

Department of Biological Sciences  
Edmonton, Alberta  
Fall 2001





University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *SHANK3*, a Candidate Gene for the 22q13 Deletion Syndrome submitted by Heather Lawrie Wilson in partial fulfillment of the requirements for the degree of Masters of Science in Molecular Biology and Genetics.



This thesis is dedicated to all families of patients with the 22q13 deletion syndrome. This study would not have been possible without your contributions.





## Abstract

Patients with the 22q13 deletion syndrome show developmental delay, delay/absence of expressive speech, hypotonia and mild dysmorphic features. The sizes of deletions in 50 cases were characterized molecularly using minisatellite, microsatellite and FISH analysis, these deletions ranged in size from microdeletions smaller than 200kb to large deletions over 10Mb. The clinical features of this syndrome were examined for correlations with the size of deletion and parent of origin of the deletion. Most features failed to show any significant correlation. Most importantly the delay/absence of expressive speech and several components of developmental delay did not show a correlation, indicating that a candidate gene for these features must be deleted in all patients. *SHANK3* was identified within the region of the NT microdeletion, and the protein is part of the cytoskeletal network of synapses joining membrane receptors to signal transduction machinery. *SHANK3* is a member of a closely related family of genes that may be responsible for proper brain development, and mutation or deletion of an allele may lead to developmental delay.





## Acknowledgements

Heather McDermid:

Thank you for being a wonderful supervisor who guided me and supported me throughout my project.

Susan Andrew, Shelagh Campbell and Paul Wong:

Thank you to the members of my committee for their suggestions and guidance throughout my project.

Katy Phelan and Gail Stapleton:

Thank your for providing me with extensive phenotype information on many of the patients in this project. And also for organizing the support group for the 22q13 deletion syndrome, which I was fortunate to attend in Aug 2000.

Polly Brinkman-Mills and Stephanie Maier:

Thank you for all of your support throughout my project and for being there to chat with. Thank you for all of your faith in my abilities.

Ranessa Cooper:

Thank you for being a great mentor and for being there for my various questions and difficulties.

Song Hu:

Thank you for all of your technical advice in helping me work out various bugs in a protocol.

Andrew Wong:

Thank you for your support in getting me started on my project.

Helen Cheung and Wai-Yue Tse:

Thank you for your help in carrying out various components of my project.

Mr. D. Smart:

Thank you for getting me interested in genetics in high school.



My Parents (Jane and Wayne Wilson):

Thank you for allowing me to work on the dining room table for a month and for all of your support.

My Sister (Kathleen Wilson):

Thank you for your support and forcing me to get to work some days.

All of the families of patients with the 22q13 deletion syndrome:

Thank you for providing the necessary samples and information to all this analysis to be carried out.





# **TABLE of CONTENTS**

## **Chapter 1: Introduction**

## **Page**

Chromosome 22	1
Cat Eye Syndrome	1
DiGeorge Syndrome/Velocardiofacial Syndrome	2
22q13 Deletion Syndrome	2
Phenotype of the 22q13 Deletion Syndrome	3
Cytogenetics of the 22q13 Deletion Syndrome	6
Molecular Characterization of the 22q13 Deletion Syndrome	11
NT 22q13 Microdeletion	13
Search for Candidate Genes	15
<i>Acrosin</i> Gene	17
<i>RABL2B</i> Gene	17
“ <i>ALPR</i> ” Gene	18
Telomeres	20
Telomere Position Effect in NT	22
Chromosome 22 Ring Chromosomes	23
Research Objectives	25

## **Chapter 2: Materials and Methods**

Selection of Patients	34
Isolation of Nucleic Acids	36
DNA Extraction from Whole Blood	36
Isolation of Genomic DNA from Cell Line	37
Isolation of Plasmid DNA from Bacterial Cells	38
Isolation of Cosmid DNA for FISH Probes	39
Isolation of Total RNA from Tissue Samples	40
Isolation of mRNA from total RNA	41
Fluorescent <i>In Situ</i> Hybridization (FISH)	42
Preparation of Metaphase Cells for use in FISH	42
Labeling FISH Probes	43
Hybridization and Detection of FISH Probes	45





Southern and Northern Blot Analysis	47
Preparation of Southern Blots	47
Preparation of Northern Blots	48
Labeling of Radioactive Probes for Northern and Southern Blots	49
Southern Blot Hybridization	50
Northern Blot Hybridization	51
Stripping Northern and Southern Blots for Re-use	52
Polymerase Chain Reaction Analysis of DNA Samples	52
Microsatellite Polymerase Chain Reaction	52
Reverse Transcription from RNA	54
Polymerase Chain Reaction to Determine Genomic Structure and Transcripts	55
Extraction of PCR Product from Agarose Gel	56
DNA Manipulations	56
Ligation of DNA into Plasmids	56
Transformation of <i>E. coli</i> Cells with Plasmid DNA	56
Sequencing DNA	57

### **Chapter 3: Results**

Characterization of the Novel Gene in the NT Deletion Region	61
Assembly of <i>SHANK3</i> Sequence	61
Characterization of <i>SHANK3</i> Transcripts	63
Expression of <i>SHANK3</i>	66
Expression of <i>Shank3</i> in Other Organisms	66
Characterization of Another Member of the Shank Family of Proteins	67
Characterization of Patient Deletions	68
Verifying the Deletion of <i>SHANK3</i> in Patients	68
Determining Deletion Size Using Microsatellites	70
Determining Deletion Size Using FISH	71
Determining Parent of Origin of Deletion	73
Analysis of the Phenotype of the 22q13 Deletion Syndrome Patients	73
Characterization of Unusual Patients	75
Patients with Bisatellited Chromosome 22	75
Patient with TAPVR	76
Translocation Patient t(14;22)	76
Mosaic Patient (JaA)	77
Interstitial Deletion (MRS)	77



## **Chapter 4: Discussion**

Characterization of the human ortholog of rat <i>Shank3</i>	121
<i>SHANK3</i> Transcript Analysis	122
Expression of <i>SHANK3</i>	127
Future Work	132
Analysis of Patient Deletions	133
Analysis of Patient Phenotype	137
Analysis of Genes in 22q13 Deletion Region	141
Future Work	146
Characterization of Unusual Patients	146
Bisatellited Chromosome 22	146
Patient with TAPVR and 22q13 Deletion	148
Translocation Patient t(14;22)	148
JaA Mosaic Deletion	150
Interstitial Deletion (MRS)	150
Future Work	153
Other Future Experiments	153
Overall Conclusions	157

## **Bibliography**

<b>Appendix A</b>	% GC Formula for Calculating $T_m$ of Primers	174
<b>Appendix B</b>	Questions used in analysis of phenotype of patients with the 22q13 deletion syndrome	175
<b>Appendix C</b>	Example of Linear Regression Analysis Output	177
<b>Appendix D</b>	Example of Contingency Table and Chi Square Analysis Output	178
<b>Appendix E</b>	Web Sites and Computer Software Used	180



## **List of Tables:**

<b>Number</b>	<b>Title</b>	<b>Page</b>
Table 1-1:	Summary of cytogenetic and clinical features of published patients with a deletion of 22q13	32-33
Table 2-1:	Microsatellites Used in Analysis of Deletions	60
Table 3-1:	PCR products obtained in genomic analysis of SHANK3	108
Table 3-2:	Summary of Minisatellite and Microsatellite Analysis	109-110
Table 3-3:	Summary of FISH Analysis	111-112
Table 3-4:	Summary of parent of origin, karyotype and molecular breakpoints of patients	113-114
Table 3-5:	Summary of clinical features of deletion patients	115-118
Table 3-6:	Summary of Statistical Analysis Comparing Size of Deletion and Clinical Features Observed	119
Table 3-7:	Summary of Statistical Analysis Comparing Parent of Origin of Deletion and Clinical Features Observed	120





## **List of Figures:**

<b>Number</b>	<b>Title</b>	<b>Page</b>
Figure 1-1:	Ideogram of Chromosome 22	27
Figure 1-2:	Diagram of cytogenetic information of published patient deletions	28
Figure 1-3:	Diagram of molecular information of published patient deletions	29
Figure 1-4:	Diagram of “ALPR” as identified by A. Wong (1998)	30-31
Figure 3-1:	Alignment of Human <i>SHANK3</i> with Rat <i>ProSAP2</i> at the nucleic acid level	79-84
Figure 3-2:	Genomic organization of SHANK3	85-86
Figure 3-3:	Similarity between SHANK3 and other Shank family members at the amino acid level	87-88
Figure 3-4:	Alignment of human SHANK3 domains with other Shank gene family members	89-92
Figure 3-5:	Northern Blot Analysis of SHANK3 in humans	93-94
Figure 3-6:	Mouse multiple tissue Northern Blot Analysis	95
Figure 3-7:	Pig Multiple Tissue Northern Blot Analysis	96-97
Figure 3-8:	Example of a result for minisatellite D22S163	98
Figure 3-9:	Example of a microsatellite gel	99
Figure 3-10:	Examples of microsatellite analysis	100
Figure 3-11:	Sample of FISH results	101-102
Figure 3-12:	Summary of sizes of patient deletions	103-104
Figure 3-13:	Patients 4’s bisatellited chromosome	105
Figure 3-14:	FISH results from patient 2	106-107



Figure 4-1:	Protein components of the post-synaptic density of excitatory synapses	158-159
Figure 4-2:	Genomic structure and predicted transcripts of <i>SHANK3</i>	160-161
Figure 4-3:	Diagram of region deleted in 22q13 deletion syndrome	162-163



## **Abbreviations:**

AE: Age Equivalent

ALPR: Ankyrin Like Proline Rich

ARSA: Arylsulfatase A

BAC: Bacterial Artificial Chromosome

BLAST: Basic Local Alignment Search Tool

BSA: Bovine Serum Albumin

CaCo: Colon Cancer cell line

CARS: Childhood Autism Rating Score

CELSR1: Cadherin EGF LAG seven-pass G-type receptor 1

CES: Cat Eye Syndrome

cM: centimorgan

CPT1B: Carnitine palmitoyltransferase I, muscle

CortBP: Cortactin Binding Protein

COX: cytochrome oxidase

DA/DAPI: stain specific for satellite of chromosome 15

der: derivative chromosome

DGS/VCFS: DiGeorge Syndrome/Velocardiofacial Syndrome

DP: Developmental Profile II

DP A: Academic Skills

DP C: Communication Skills

DP IQ: Intelligence Score





DP P: Physical Skills

DP S: Social Skills

DP SH: Self Help Skills

ECGF1: Endothelial cell growth factor 1 (platelet-derived)

EST: Expressed Sequence Tag

FISH: Fluorescent *In Situ* Hybridization

FITC: Fluorescein label on antibody

GFP: Green Fluorescent Protein

GKAP: Guanylate Kinase-Associated Protein

HeLa: cancer cell line

IPTG: Isopropyl Thiogalactoside

kb: kilobase pairs

LB: Luria-Bertani medium

MAPK: Mitogen Activated Protein Kinase

Mb: megabase pairs

MLC: megalencephalic leukoencephalopathy

MOV10L1: Mov10 (mouse)-like 1

NBT: nitroblue tetrazolium

NCBI: National Center for Biotechnology Information

NMDA Receptors: N-methyl-D-aspartate receptors

NOR: Nucleolar Organizing Region

OMIM: Online Mendelian Inheritance in Man ([www.ncbi.nlm.nih.gov/Omim/](http://www.ncbi.nlm.nih.gov/Omim/))

ORF: Open Reading Frame



PAC: P1 Artificial Chromosome

PCR: Polymerase Chain Reaction

PDZ: PSD-95, Dlg, ZO-1

PHA: phytohemagglutinin

PKDREJ: Polycystic kidney disease (polycystin) and REJ (sperm receptor for egg jelly, sea urchin homolog)-like

ProSAP: Proline rich Synapse Associated Protein

PSD: Postsynaptic Density

r(22): ring chromosome of 22

RABL: RAB, member of RAS oncogene family-like

RACE: Random Amplification of cDNA Ends

RFLP: Restriction Fragment Length Polymorphism

RNAi: RNA interference

RT: Reverse Transcriptase

SAGE: Serial Analysis of Gene Expression

SAM: Sterile Alpha Motif

SAP: Synapse Associated Protein

SAPAP: SAP Associated Protein

SA-PMP: Streptavidin-paramagnetic particles

SCO2: SCO (cytochrome oxidase deficient, yeast) homolog 2

SH3: Src homology 3

SHANK: SH3 Ankyrin repeat

SIB: Scales of Independent Behavior



SIB BI: Broad Independence

SIB CL: Community Living Skills

SIB MS: Motor Skills

SIB PL: Personal Living Skills

SIB S: Support

SIB SC: Social and Communication Skills

SPANK: SH3 domain, PDZ domain, Ankyrin Repeats

SSTRIP: Somatostatin Receptor Interacting Protein

t( ): translocation

TAPVR: Totally Anomalous Pulmonary Venous Return

VNTR: Variable Number of Tandem Repeats

VSD: Ventricular Septal Defect

X-Gal: 5-bromo-4-chloro-3-indol- $\beta$ -D-galactosidase





## **Chapter 1: Introduction**

### **Chromosome 22:**

Chromosome 22 is the second smallest autosome in the human genome and contains approximately 1.6% of total genomic DNA (figure 1-1). Due to its small size chromosome 22 was the first chromosome to be completely sequenced (Dunham, et al, 1999). This chromosome is one of the 5 acrocentric chromosomes found in a human karyotype. The short arms of all the acrocentric chromosomes, including chromosome 22, are similar in sequence and structure. The small p-arms contain highly repetitive DNA composed of tandemly repeated ribosomal RNA genes as well as other tandem repeat arrays. Several well characterized genetic disorders have been mapped to chromosome 22 and have driven the analysis of this chromosome. Two of the best characterized syndromes are Cat Eye syndrome (CES) and DiGeorge/Velo-Cardio-Facial syndrome (DGS/VCFS), both of which localize to the q arm.

### **Cat Eye Syndrome (CES):**

Cat Eye syndrome is the result of a duplication of 22q11, usually in the form of a supernumerary chromosome that contains an inverse duplication of part of 22q11 (McDermid, et al, 1986). CES was named since one of the clinical features of this syndrome is ocular coloboma, particularly of the iris, giving these patients eyes that appear cat-like (Schinzel, et al, 1981; OMIM: 115470). Several other clinical features are commonly associated with CES including: heart malformations (such as totally anomalous pulmonary venous return (TAPVR) and tetralogy of Fallot), kidney malformations (such as agenesis, and horseshoe shaped kidney), anal atresia,



downslanting palpebral fissures, and preauricular pits and tags. CES patients do not frequently show developmental delay, if a developmental delay is found in a CES patient then it is relatively mild.

#### DiGeorge/VCF Syndrome (DGS/VCFS):

DiGeorge/VCF syndrome results from the deletion of 22q11. DGS/VCFS is the best characterized disorder associated with the deletion of chromosome 22 and has been called the 22q deletion syndrome (Greenberg, et al, 1988; Carlson, et al, 1997). The key clinical features associated with DGS/VCFS are outflow tract defects of the heart (such as tetralogy of Fallot, truncus arteriosus, ventricular septal defect (VSD), and interrupted aortic arch) and several facial dysmorphisms including: telecanthus, low set ears, small mouth, short philtrum, bulbous nose and palatal clefting (OMIM: 18840, 192430; Shprintzen, et al, 1981). Patients with DGS/VCFS have mild to moderate learning difficulties and may show psychiatric disorders, such as schizophrenia. The breakpoints for DGS/VCFS have been found to occur within low copy repeats that are 97 to 98% identical (Morrow, et al, 1995; Carlson, et al, 1997; Shaikh, et al, 2000). DGS/VCFS is frequently caused by a microdeletion of 22q11 that is not cytogenetically visible. Therefore FISH probes with a control probe on 22q13 were developed by companies to test for DGS/VCFS.

#### 22q13 Deletion Syndrome:

A second less frequent deletion syndrome is found at the terminal end of chromosome 22q. The first confirmed case of the 22q13 deletion syndrome was published in 1985 (Watt, et al). This patient was identified through cytogenetics due to a



rearrangement of a familial pericentric inversion of chromosome 22 that resulted in partial monosomy for chromosome 22. Many other cases of the 22q13 deletion syndrome have since been published.

Approximately 35 patients with the 22q13 deletion syndrome have been identified in the literature or abstracts (table 1-1). Many of the reports were of single patients or small groups of two or three patients identified with a deletion of the terminal end of chromosome 22. In most of these reports the patients were identified as having the 22q13 deletion syndrome but the sizes of the deletions were not characterized molecularly. Many of the patients were identified as having a cytogenetically visible deletion, or another chromosomal abnormality, such as a translocation, that would allow for cytogenetic detection of an abnormality with chromosome 22. Since chromosome 22 is a predominantly G-light chromosome, especially in the terminal band, it is difficult to cytogenetically identify small terminal deletions.

#### Phenotype of the 22q13 Deletion Syndrome:

A recent review of the 22q13 deletion syndrome studied the phenotype of 37 patients and compared the phenotype to other cases presented in the literature (Phelan, et al, 2001). The individuals that participated in this study ranged in age from 12 months to 26 years with an approximately equal distribution of male and females. All 37 cases presented with global developmental delay and absence/severely delayed expressive speech. Hypotonia was present in 97% of these cases, and 95% showed normal to accelerated growth. Several other phenotypic features were noted including: increased tolerance to pain (86%), abnormal toenail growth (78%), chewing behavior (70%), large





fleshy hands (68%), dysplastic ears (65%), pointed chin (62%), dolicocephaly (57%), ptosis (57%), tendency to overheat (51%), epicanthal folds (41%), 2/3 syndactyly (38%), seizures (27%), and clinodactyly (14%). Phelan et al, noted that several of these clinical features were not widely reported in the literature. High pain tolerance had been previously reported in two cases in the literature (Precht, et al, 1998, and Goizet, et al, 2000). Abnormal toenail growth was only reported in two of the cases in the literature (Smith, et al, 1996, and Schroder, et al, 1998). No cases in the literature described patients as having large fleshy hands (large hands were described by Schroder, et al, 1998, and Nesslinger, et al, 1994, without reference to if the hands were fleshy), or having a tendency to overheat. The key clinical features of the 22q13 deletion syndrome are developmental delay, delay/absence of expressive speech and mild dysmorphic features.

Many of the patients participating in the study by Phelan, et al (2001) showed features typical of autism. The Childhood Autism Rating Scale (CARS) was used to test 18 of the 37 patients. Of the patients tested with CARS, 12 were classified in the moderate-severe autistic range and 5 were classified in the mild range. However, the authors noted that the low scores on the test for autism may be directly correlated to the fact that these children are severe to profoundly developmentally delayed and have low adaptive behavior skills. Therefore patients with the 22q13 deletion syndrome, may show some typical clinical features of autism, but not be true cases of autism. Prasad, et al (2000) described a patient that has severe autism and believed that autism is a component of the 22q13 deletion syndrome phenotype. Goizet, et al (2000), also described a patient



who was characterized with a moderate autistic disorder. Therefore, autistic features of the 22q13 deletion syndrome may not be true autism but may be the result of the other clinical features including the severe to profound developmental delay and the delay/absence of expressive speech.

Very few cases of the 22q13 deletion syndrome have been identified prenatally, the majority of the cases were diagnosed in early childhood. In the study by Phelan, et al (2001), only two cases were detected prenatally with the amniocentesis being performed due to advanced maternal age. During the first two years of life 17 out of 33 cases were diagnosed as having the 22q13 deletion syndrome. The oldest age at diagnosis was 20 years of age. A recent abstract identified an individual with the 22q13 deletion syndrome that is 46 years of age (Cassidy, et al, 1999). This particular individual is of interest since none of the other patients with a known 22q13 deletion have reached this age. Little information is available on this patient, but characterization of the course of development of this genetic syndrome in this individual may allow for predictions about the fate of other individuals with a similar deletion.

Diagnosis of the 22q13 deletion syndrome is difficult. First, only 50 cases of this syndrome have been published in the literature, making it an uncommon syndrome and one that is not well known by clinicians. Second, the clinical features associated with this syndrome are relatively non-specific and may resemble other syndromes. These patients show developmental delay and delay or absence of expressive speech, both of which are common to many genetic disorders. The dysmorphic features commonly observed in patients with the 22q13 deletion syndrome are mild and no particular feature is indicative



of a deletion of 22q13. If a diagnosis of the 22q13 deletion syndrome is made based on the phenotype then it must be confirmed using cytogenetics.

#### Cytogenetics of the 22q13 Deletion Syndrome:

The first case of the 22q13 deletion syndrome (Watt, et al, 1985) was identified based on an unusual chromosome 22 in the karyotype. This chromosome was identified as having satellite material at both ends (part of the p-arm of the chromosome was duplicated) and a deletion of the terminal end from 22q12 to qter (figure 1-2).

Karyotypes of family members showed that the mother, brother and sister all carried a chromosome 22 containing a pericentric inversion (inv(22)(p11q12)). The patient's unusual chromosome is believed to have arisen from a recombination event involving this pericentric inversion. Such an unusual chromosomal rearrangement in this patient allowed for the identification of the deletion of 22q13.

Several cases of the 22q13 deletion syndrome were reported in the late 1980's and early 1990's. These cases were all identified on the basis of cytogenetically detectable deletions. The patient published in Herman, et al (1988) showed a de novo terminal deletion with the breakpoint located in 22q13.31. Kirshenbaum, et al (1988) also identified a patient with a terminal deletion of chromosome 22q with the breakpoint located in 22q12; representing a larger deletion than the patient in Herman, et al (1988). Several cases have since been identified using high resolution GTG banding of chromosomes. However, since the terminal region of chromosome 22q stains very lightly in conventional G-banding, detection of subtle terminal deletions (less than approximately 5Mb) is difficult unless there is another chromosomal abnormality such as





an inversion or a translocation.

Two reports in the early 1990's identified individuals as having partial monosomy of 22q13 as the result of a translocation. The first case was identified as having an unbalanced karyotype resulting from the segregation of a paternal balanced translocation (Zackowski, et al, 1990). The patient's karyotype was cytogenetically determined to be 46,XX,-22, +der(22)t(7;22)(q32;q13.3). This means that this patient was monosomic for 22q13.3 and trisomic for 7q32. Several clinical features of the 22q13 deletion syndrome were identified in this patient but the patient was too young to ascertain the developmental delay and speech ability. The second case was also identified as an unbalanced karyotype resulting from the segregation of a maternal balanced translocation (Tachdjian, et al, 1992). The patient's karyotype was determined to be 46,XX,-22, +der(22)(t11;22)(q23.3;q13.2). This means that this patient was monosomic for 22q13 and trisomic for 11q23.3. Several clinical features of the 22q13 deletion syndrome were characterized in this patient but the patient died at 10 months. Unbalanced translocations that result in the deletion of 22q13 can make identification of deletions cytogenetically easier but can complicate the study of the phenotype since two chromosomal regions are involved and both may contribute to the phenotype.

Mosaicism of a chromosomal abnormality can also confound the relationship of the abnormality to the phenotype observed. The first patient identified as mosaic for the 22q13 deletion was described by Yong, et al. (1997). Both developmental delay and absence of expressive speech were observed in this patient. Several other clinical features commonly associated with the 22q13 deletion syndrome such as prominent ears,



long philtrum, and febrile convulsions were observed in the patient. Cytogenetically this patient was found to have a deletion of one chromosome 22 at 22q13.2 in some of the cells examined. Further FISH analysis was carried out using the DGS/VCFS deletion probe and its control probe from 22q13.3 (ARSA, figure 1-3). Every cell examined showed the DGS/VCFS probe on both chromosome 22s, but only 1 out of 15 cells examined showed two signals from the control probe on both of the chromosome 22s. The loss of genetic material from a normal cell is more probable than the gain of material in an abnormal cell, and therefore the fertilized egg that produced this individual likely contained a normal chromosome complement, and the deletion occurred in the zygote. Since the majority of the cells are deleted for 22q13, the deletion likely occurred quite early in development. This was the first case of mosaicism for the deletion of chromosome 22 that is not due to a ring chromosome.

Very few cases of the 22q13 deletion syndrome have been identified prenatally. One case of prenatal diagnosis was published by Riegel, et al (2000). In this case, a cystic tumor on the neck of the fetus was detected by ultrasound at 21 weeks and further analysis was carried out. Karyotyping from a chorionic villus sampling showed a deletion of the long arm of chromosome 22 in 10 out of 11 cells. FISH analysis was carried out using the DGS/VCFS probe and its control probe from 22q13.3 (ARSA, figure 1-3). The fetus was aborted at 23 weeks and only mild dysmorphisms were noted at the autopsy. The clinical features of the 22q13 deletion syndrome would not be visible on an ultrasound; therefore prenatal diagnosis of the 22q13 deletion syndrome is difficult unless a large deletion or a translocation is present in a karyotype conducted for other



reasons.

The review by Phelan, et al (2001) compiled the cytogenetic information on the 37 patients in their study. Most of the cases were terminal deletions (29/37), while the remainder of the patients had unbalanced translocations (8/37). Two of the individuals identified with terminal deletions were found to be mosaic for this deletion; one of the individuals had an additional cell line with a r(22) chromosome and the other individual had an additional cell line with a normal chromosome complement. Failure to diagnose the 22q13 deletion syndrome in previous cytogenetic studies occurred in 12 of the individuals in this study, including three amniocenteses, which illustrates that the 22q13 deletion syndrome is difficult to confirm cytogenetically. This comes from the fact that the terminal region of 22q is a G-light region that is very pale, with almost no cytogenetic landmarks. Many deletions of 22q13 are also quite small and are difficult to visualize using conventional cytogenetics, and thus another technique must be used.

Fluorescent in situ hybridization (FISH) has been used to detect many subtle deletions of 22q13. The first case of a deletion of 22q13 that was undetectable cytogenetically but was identified by FISH analysis was the result of a cryptic translocation (Smith, et al, 1996). This patient was karyotyped and a subtle chromosomal abnormality was detected on chromosome 17. Using FISH with a probe specific to chromosome 17, signals were detected on both chromosome 17s as well as an acrocentric chromosome. Using one of the control probes for DGS/VCFS (D22S39), a deletion of 22q13 was identified (figure 1-3). Two other cryptic translocations involving chromosome 22 were identified by Doheny, et al (1997). The first case was suspected to



have an abnormality of the terminal band of chromosome 22q from results of routine cytogenetics. The probe D22S39 was used on metaphase spreads from this patient and verified that there was a terminal deletion of 22q13. Analysis of the mother's chromosomes confirmed that the deletion in the child was due to an unbalanced karyotype resulting from a maternal balanced translocation with chromosome 1q. The second case was also suspected to have an abnormality of the terminal band of chromosome 22 based on cytogenetic results. A deletion of the FISH probe D22S39 was detected in the metaphase spreads from this patient. A translocation of 6p to the terminal region of 22q was later characterized in this patient (Praphanphoj, et al, 2000). All three of these cases showed some indication of a chromosomal abnormality based on cytogenetic results.

Several cases of the 22q13 deletion syndrome were originally clinically characterized as another genetic disorder but FISH analysis showed that they were due to the deletion of 22q13. The two patients reported in Precht, et al (1998), were referred for FISH analysis to characterize the cause of the clinical features. Patient one was referred for FISH analysis to eliminate the diagnosis of Angelman syndrome, which is caused by a deletion of chromosome 15. This patient showed some of the clinical features of Angelman syndrome including absence of expressive speech, hypotonia and an unsteady gait. FISH analysis was negative for a deletion of the Angelman/Prader-Willi probe. By chance, this patient was used as a control patient for FISH analysis of a DGS/VCFS patient. Unexpectedly the control probe for the DGS/VCFS test (Oncor), D22S39, was deleted in this individual (figure 1-3). The second patient was referred for FISH analysis





to verify the clinical diagnosis of DGS/VCFS. The patient was described as having facial features that were similar to those observed in DGS/VCFS. When the FISH analysis was carried out on this patient the control probe D22S39 was found to be deleted. Two of the patients published by Prasad, et al (2000) were originally referred for FISH analysis to verify the clinical diagnosis of DGS/VCFS. Using the DGS/VCFS probe from Vysis Inc., the control probe (the Arylsulfatase A gene, ARSA) at the terminal region was noted to be absent in the first patient. The second patient was analyzed using the DGS/VCFS test from Oncor and a deletion of the control probe D22S39 was found. The patient identified by DeVries, et al (2000), was originally clinically characterized as having FG syndrome. FG is an X-linked recessive disorder that includes developmental delay along with hypotonia and mild facial dysmorphisms. FISH analysis with a set of telomeric probes identified this patient as having a terminal deletion of 22q. The clinical characterization of the 22q13 deletion syndrome is a difficult task, but with the use of FISH analysis more definitive diagnosis can be made.

#### Molecular Characterization of the 22q13 Deletion Syndrome:

Many of the cases of the 22q13 deletion syndrome are simply identified as having the 22q13 deletion syndrome based on cytogenetics or a few FISH probes. The first study that characterized a deletion molecularly was Phelan, et al (1992). The patient in this study, FB, was cytogenetically identified as having a terminal deletion of 22q. VNTRs were used to determine the parent of origin and showed that this patient had a paternally derived deletion that included the ARSA gene (figure 1-3). One study in 1998 (Precht, et al) characterized the size and parent of origin of deletions in individuals using



microsatellite analysis. Five microsatellite markers were selected in the terminal region. Both patients showed a paternal deletion of the most distal microsatellite marker, whereas the four more proximal microsatellite markers showed no deletion. Another group in 1998 (Schroder, et al) used microsatellite and FISH analysis to further refine the size of deletions in three patients. The smallest deletion characterized in this study was approximately 26 cM in size. All of the studies where the patients were characterized at a molecular level only characterized a few patients.

The most comprehensive report on the molecular aspects of the 22q13 deletion syndrome identified seven patients and described the phenotype and the size of the deletion present in these patients (Nesslinger, et al, 1994). All seven patients had hypotonia, developmental delay, delay of gross motor milestones, and delay of expressive speech. Of the other clinical features noted in these patients, some were common to many of the patients, whereas others were specific to only one patient. All of the patients in this report have cytogenetically identifiable deletions of 22q13.3. The sizes of these deletions were determined using RFLP and Southern dosage analysis (figure 1-3). When this study was carried out the most distal probe known on chromosome 22 was ARSA, which was deleted in all seven of these patients. The proximal breakpoints in these seven patients were found to vary over a distance of 13.8cM. When the size of the deletion was compared to the various clinical features observed, no obvious correlations were detected. However, only seven patients were characterized in this study and a small variation in the observed clinical features could skew the analysis. The size of the smallest deletion examined in the Nesslinger study was estimated to be about 5Mb in size as measured by



pulsed field gel electrophoresis (McDermid, unpublished data). The critical region for the clinical features observed in the 22q13 deletion syndrome can be defined by the size of the smallest deletion. This means that the critical region for the 22q13 deletion syndrome in the Nesslinger study is 5Mb in size. A critical region of 5Mb is quite large, and too big to make the search for candidate genes feasible.

In 1995, Flint et al carried out a study looking for cryptic telomeric chromosomal abnormalities in patients with idiopathic mental retardation in hopes of identifying potential causes of mental retardation. In this study 28 terminal regions were examined on 99 patients with mental retardation using highly polymorphic probes. Three patients in this sample were identified as having a terminal deletion: two had a deletion of 22q and one had a deletion of 13q. The first 22q deletion, LM, was due to a cryptic de novo translocation to chromosome 9q. LM shows the typical clinical features associated with the 22q13 deletion syndrome. Using the minisatellite D22S163, a maternal deletion was detected. The second patient with a terminal deletion of 22q, NT, was later identified as a microdeletion, and remains the smallest deletion associated with this syndrome.

### **NT 22q13 Microdeletion:**

The patient NT (Flint, et al, 1995) has mild developmental delays, delay of expressive speech, and bilateral accessory nipples. NT was detected by a paternal deletion of the probe D22S163. Several more probes were used to determine the proximal breakpoint in NT, but none of these probes showed a deletion, indicating that this patient has a small terminal deletion of 22q13.3.

Since patient NT had such a small deletion and had two of the key features of the



22q13 deletion syndrome, it was now possible to start the search for candidate genes for these features. Wong, et al in 1997 carried out molecular characterization of the NT microdeletion. Using pulsed field gel electrophoresis, the probe D22S163 was located within 120kb of the gene ARSA. A cosmid walk of two cosmids was then carried out both proximally and distally of the probe D22S163. The proximal cosmid, cN85E7, was found not to be deleted in NT by densitometric and RFLP analysis. The distal cosmid, cN66C4, was found to be deleted in NT. A further cosmid walk in the distal direction from cN66C4 identified three more overlapping cosmids between cN66C4 and the telomere that span approximately 100kb. FISH analysis was carried out using the most telomeric cosmid to verify that the deletion in NT was a terminal deletion.

By Southern blot analysis, the breakpoint of the NT deletion was identified to be within the probe D22S163. PCR analysis was used to amplify the NT breakpoint and the product was sequenced. The breakpoint was localized to the 3' end of the D22S163 minisatellite and was healed by the addition of telomeric repeats (ttaggg). The gene *Acrosin* was localized to the NT microdeletion region using a Southern blot of the cosmids probed with a cDNA probe.

The cosmids identified by Wong, et al (1997) formed part of the effort to sequence chromosome 22. These cosmids were submitted to the University of Oklahoma Department of Chemistry and Biochemistry for sequencing (Dunham, et al, 1999). The sequence of these cosmids is available in the public databases (accession numbers: AC000050, AC000036, AC002050, AC002056). With the sequence of the cosmids available, the region could be examined for candidate genes for the neurological





abnormalities seen in patient NT.

### **Search for Candidate Genes:**

A candidate gene that causes at least some of the clinical features observed with the 22q13 deletion syndrome would have to be dosage sensitive. Since this is a deletion disorder, all individuals with this disorder would have a single copy of any candidate genes. Haploinsufficiency is found in many different genetic disorders and is frequently associated with deletion syndromes. In order for a gene to show haploinsufficiency, it must be a gene whose transcription is tightly regulated and whose products must be in particular balance with other gene products (Fisher, and Scambler, 1994). Types of genes that would be sensitive to haploinsufficiency include: 1) genes that are a part of a signaling system such as receptors and various signal transduction molecules; 2) genes that encode a structural protein that interacts with other proteins and must be in specific proportions to ensure that the final structure is correct; 3) genes that are transcriptional regulators. Since these proteins commonly compete for promoter sites, a slight imbalance in the ratios would affect the ability of the protein to properly function.

An example of haploinsufficiency in a genetic disorder is Williams Syndrome caused by a deletion of 7q11.23. Two genes in this region were identified as dosage sensitive genes. The first gene characterized to be hemizygous in William's syndrome is the *elastin* gene (Ewart, et al, 1993). The elastin protein plays an important role in the elasticity of fibers. Patients that are hemizygous for the elastin protein only have half the amount available to their cells. Due to the decrease in elastin the large arteries of the vascular system do not have the elasticity to function; these vessels become weakened



resulting in supraaortic stenosis (Ewart, et al, 1994). Several other tissues are affected by the loss of elastin and all are part of the connective tissues. The *LIMK1* gene is also found in the William's syndrome deletion region and encodes a protein predicted to be part of an intracellular signaling pathway (Frangiskakis, et al, 1996; Monaco, 1996; Tassabehji, et al, 1996). This intracellular signaling pathway may be part of the neural network, and the loss of a part of the pathway has been hypothesized to contribute to neurological abnormalities observed in Williams syndrome. Another genetic disorder that shows dosage sensitivity is Aniridia that is caused by the deletion or mutation of *PAX6*, a transcriptional regulator, located at 11p13 (Ton, et al, 1991). Several other genetic disorders have been identified as showing haploinsufficiency, but no candidate gene has been identified for many of these genetic diseases.

In order to identify a candidate gene for the 22q13 deletion syndrome, a critical region for causing the phenotype needs to be identified. To determine a critical region, the sizes and location of all characterized deletions need to be compared. The critical region is the smallest region that is deleted in all patients identified with similar clinical features. All patients published in the literature have developmental delay and delay or absence of expressive speech. Since both of these phenotypes are found in the NT microdeletion, we used patient NT's deletion to define the critical region for the neurological abnormalities associated with the 22q13 deletion syndrome. The deletion of NT is only 130kb in size, small enough to be examined for candidate genes. Only three genes were identified to be deleted in the NT microdeletion.



### Acrosin Gene:

The first gene identified in this region of the NT microdeletion is *Acrosin* (Wong, et al, 1997). *Acrosin* encodes a proteinase that is localized to the acrosome of mature spermatozoa (OMIM: 102480). *Acrosin*, when active, lyses the zona pellucida of the egg to allow for penetration of the sperm into the ovum. Since the *Acrosin* gene has a very specific function in fertilization, it is unlikely that this gene would be responsible for any of the neurological phenotypes observed in the 22q13 deletion syndrome. The only effect that deletion of *Acrosin* may have on the individuals with the 22q13 deletion syndrome would be reduced male fertility (OMIM: 102480). However, since all of the patients with the 22q13 deletion syndrome are severe to profoundly developmentally delayed, the likelihood of these individuals reproducing is very low. Thus *Acrosin* can be discarded as a good candidate gene for the neurological abnormalities observed with this syndrome.

### RABL2B Gene:

A second gene identified in this region, was *RABL2B* (Wong, et al, 1999). *RABL2B* is part of the RAB family of GTPases involved in vesicular trafficking in cells. *RABL2B* is the most telomeric gene on chromosome 22q and is located adjacent to the sub-telomeric repeats (about 2.5kb proximal of repeats). FISH analysis with the most telomeric cosmid that contains *RABL2B* showed signals on the telomere of 22q as well as near the centromere of chromosome 2q. Weaker signal with this probe was also detected on several different chromosomal ends, probably caused by the adjacent repeats of the subtelomeric region. A second locus of *RABL2*, *RABL2A*, was identified on chromosome 2. Comparison of the sequence of *RABL2A* and *RABL2B* showed only five



nucleotide differences. Northern blot analysis of these two genes detected several transcripts in most of the adult and fetal tissues tested. RT-PCR analysis distinguishing the transcripts from the two genes showed that both genes are actively transcribed. Since four copies of *RABL2* are actively transcribed, the loss of the single copy (as in the patients with the 22q13 deletion syndrome) is not likely to have a significant effect on the function of the cells.

#### “ALPR” Gene:

A third novel gene was identified within the region of the NT microdeletion by A. Wong (1998). *Ankyrin-Like Proline Rich (ALPR)* was identified by predicting two groups of exons in the NT deletion region using the GRAIL computer program. *ALPR* would eventually go through several name changes (*PROSAP2*, *Spank2*) before becoming known as *SHANK3*. When an ORF was predicted from these two groups of exons, it matched a predicted protein in *C. elegans* C33B4.3. The predicted ORF also showed homology to the rat cortactin binding protein. *ALPR* was partially cloned by A. Wong using a combination of RT-PCR, cDNA clones from library, EST contigs, and prediction programs (figure 1-4). Five different partial cDNAs of this gene were identified but no full length cDNA was found (figure 1-4). Several exons of this gene were only predicted, and not found in any cDNA. A. Wong identified three alternative 3' ends to this gene indicating that there are several transcripts produced from *ALPR*.

Partial cDNAs of *ALPR* were used to probe Northern Blots to obtain an idea of the expression pattern (Wong, 1998). Different sized bands were observed in brain tissue as well as several of the other tissues tested. A transcript of approximately 7.5kb was





observed in adult heart, brain and placenta, and weaker expression in other tissues, with two of the probes, FL2 and 55337F-R1. A transcript of 6.8kb was observed only in adult brain using these two probes. On fetal tissue blots three transcripts were detected with these two probes: a large >10kb transcript, a transcript about 8kb and another of about 7.5kb detected in fetal brain, kidney and lung. A. Wong hypothesized that one of the alternative 3' ends FLS was specific to fetal development since no transcript was visible on adult tissues when probed with this exon. FLS identified a >10kb large transcript and a transcript of 2.5kb that was specific to fetal liver. Northern blot analysis of *ALPR* identified a brain specific transcript, thus making *ALPR* a potential candidate gene for the neurological abnormalities observed with the 22q13 deletion syndrome.

A. Wong (1998) hypothesized about the function of this novel gene, *ALPR*. Since *ALPR* has homology to the rat cortactin binding protein1 (CortBP1) the function was hypothesized to be involved in cytoskeleton regulation in a similar manner to CortBP1. The ankyrin repeats identified at the 5' end of *ALPR* indicated that the protein may be membrane bound or may be involved in protein-protein interactions. The proline rich region of *ALPR* was hypothesized to be involved in protein-protein interactions. Since there are several transcripts identified for *ALPR* in many different tissues, *ALPR* may have several different functions depending on the transcript produced. Little else is known about the potential function of *ALPR*.

A recently identified patient with a balanced translocation between 22q13 and 12q24 has provided further evidence of *ALPR*'s involvement in the 22q13 deletion syndrome (Bonaglia, et al, 2001). This patient shows the typical clinical features



associated with the 22q13 deletion syndrome including: mild developmental delay, delay of expressive speech, hypotonia, dolicocephaly, epicanthal folds and bulbous nose. FISH analysis and inverse PCR were used to determine that the breakpoint on chromosome 22 disrupts the *ALPR* gene (the authors use the name *ProSAP2* for this gene). This patient confirms that the disruption of this gene is responsible for the neurological abnormalities observed in the 22q13 deletion syndrome.

### **Telomeres:**

Patient NT was identified through a screen of telomeres for subtle chromosome rearrangements that might be causing idiopathic developmental delay (Flint, et al, 1995). Thus it is believed that subtle chromosomal rearrangements of the telomeres may be responsible for idiopathic development delay either through the deletion or disruption of a gene, or position effect of telomeres. All of the patients with the 22q13 deletion syndrome are believed to have a terminal deletion that is either due to a translocation or a de novo deletion that is healed by the addition of telomeric repeats.

Telomeres are an important part of the chromosome as they are responsible for maintaining the correct length of the chromosome and ensuring that no genetic material is lost during replication. All human telomeres are made up of a repeat of TTAGGG that lies adjacent to a subtelomeric region which are composed of complex repetitive DNA (Flint, et al, 1997a; Flint, et al, 1997b). Since there is considerable similarity between the telomeric and subtelomeric regions of different chromosomes, frequent chromosomal rearrangement can occur in these regions (Ledbetter, et al, 1992). The region adjacent to a telomere has been found to have a high concentration of genes (Saccone, et al, 1992)



and frequent chromosomal rearrangements of these regions may result in the loss, gain or disruption of genes. Thus translocations or other chromosomal abnormalities that involve terminal regions would affect these genes and may be responsible for a variety of clinical features. One clinical feature commonly observed with terminal aberrations of chromosomes is developmental delay (Ledbetter, 1992; Flint, et al, 1995) .

Developmental delay is found in approximately 3% of the population, but many of the cases of developmental delay are idiopathic. Since it is hypothesized that deletions or rearrangements of the subtelomeric region may be responsible for many cases of idiopathic developmental delay, there has been pressure to develop a way to detect these rearrangements. The first study of telomeres and developmental delay was the study by Flint, et al in 1995 where a population of 99 individuals with developmental delay was screened for rearrangements of 28 chromosome ends. This study identified three individuals with such subtelomeric rearrangements. Several other studies have been carried out, examining patients with moderate to severe developmental delay for chromosomal rearrangements (National Institute of Health and Institute of Molecular Medicine Collaboration, 1996; Knight, et al, 1997; Ghaffari, et al, 1998; Knight, et al, 1999; Slavotinek, et al, 1999; Ballif, et al, 2000; Knight, et al, 2000a; Knight, et al, 2000b; Knight, et al, 2000c; deVries, et al, 2001). The largest of these studies (Knight, et al, 1999) analyzed the telomeres of 374 patients with unexplained developmental delay. This study found that in cases of moderate to severe idiopathic developmental delay, about 7% of the patients had a subtle subtelomeric rearrangements. All of the cases identified as having a subtelomeric rearrangements also had some physical abnormalities



associated with the developmental delay. Cases of mild idiopathic developmental delay were also found to have some subtelomeric rearrangements, but the prevalence is much lower, about 0.5%. Since telomeres are important for chromosomal function and the region adjacent to the telomeres contain many genes, screening telomeres for chromosomal rearrangements, especially in cases of developmental delay, would be useful in diagnosis.

### **Telomere Position Effect in NT:**

One might postulate that since the breakpoint in the NT deletion is healed by the addition of telomeric repeats to the broken chromosome end, a telomere position effect may affect the replication and expression of more proximal genes. In yeast it has been shown that genes that are in close proximity to the telomeres have retarded activation of the origins of replication (Ferguson, et al, 1992). In order to determine if this was the case, Ofir, et al (1999) analyzed the replication timing and the transcriptional activation of genes located in close proximity to the telomere healed break in the patient NT. A change in the replication timing of the region adjacent to the NT breakpoint was noted that shifted replication from the middle of S-phase to late S-phase. The influence of the change in replication timing was detected at a distance of 40kb away from the breakpoint. Using a DNaseI sensitivity assay the change in replication timing observed in NT was observed to be independent of the chromatin structure in this region. No change in the methylation status of the region proximal to the NT breakpoint was detected using methylation sensitive enzymes. The effect of the telomere on the transcriptional regulation of genes was also examined in NT. The closest gene proximal to the NT





breakpoint is ARSA, located at approximately 54kb from the breakpoint. Taking advantage of polymorphic restriction sites located within the transcript of ARSA, the authors were able to determine which allele(s) of ARSA were expressed in the NT microdeletion. Using RT-PCR analysis of one of the polymorphic sites both alleles of ARSA were shown to be expressed in this patient. This study indicates that the genes deleted in NT are responsible for the clinical features rather than a telomeric position effect on more proximal genes.

### **Chromosome 22 Ring Chromosomes:**

Ring chromosomes of chromosome 22, r(22), have been identified and have a closely related phenotype to the 22q13 deletion syndrome (reviewed in Hunter, et al, 1977). Ring chromosomes result from breaks in both the p-arm and the q-arm of a chromosome that heals itself by joining the two broken ends together. Ring chromosomes are notoriously unstable and may be lost during a cell division, or two ring chromosomes may crossover at meiosis and produce a double ring chromosome. Patients that have a ring chromosome share several clinical features with patients with a terminal deletion of chromosome 22q. Developmental delay is found in 90% of patients with a ring chromosome (Naritomi and Hirayama, 1988). Several other clinical features that are observed less frequently include: hypotonia (75%), microcephaly (41%), epicanthal folds (86%), full eyebrows (100%), large ears (72.7%), abnormal auricle (87%), maxillary hypoplasia, long eyelashes (71%), thick lips (77%), 2/3 syndactyly (60%), hand abnormalities, and ataxia/unsteady gait (87%). Autism has also been described in several r(22) patients (MacLean, et al, 2000). However, since these patients are severely



developmentally delayed, the test for autism may not be accurate. Patients with a ring chromosome show a large variation in the phenotype.

The variation observed in the phenotypes of patients with ring chromosomes may be the result of the instability of ring chromosomes and the ease by which mosaic individuals are produced. The extent of mosaicism in individuals may have an impact on the severity of the phenotype. The phenotype of an individual with a large number of cells that are missing the ring chromosome would likely be more severe than the phenotype of an individual with a lower percentage of cells that have lost the ring chromosome. The other reason for the phenotype variability observed with the ring chromosomes is the variability in the breakpoints. The loss of material from the p-arm of any of the acrocentric chromosomes has no impact on the phenotype. Since these chromosome arms contain highly repetitive DNA sequences that are similar to all of the other acrocentric chromosomes, the loss of a few copies will not have a significant effect. Many of the breakpoints on the q-arm are located within the 22q13 band and as a result material from the terminal region of the long arm is lost. The material that is lost in the formation of ring chromosomes is similar to the region of DNA that is lost in the 22q13 deletion syndrome and hence there is overlap of the phenotypic features observed in both of these chromosomal abnormalities.



## **Research Objectives:**

1. To characterize the novel gene *SHANK3* identified by A. Wong (1998) as a potential candidate gene for the 22q13 deletion syndrome neurological abnormalities. Little was known about this gene's function and how it might be involved in the clinical features of this syndrome. Brain expression and evidence of a variety of protein interacting domains makes *SHANK3* a good candidate gene for the neurological abnormalities observed in the 22q13 deletion syndrome, and thus is the focus of this thesis.
2. To examine the deletions in patients with the 22q13 deletion syndrome. In order for the novel gene, *SHANK3*, to be a candidate gene for the neurological abnormalities, it needs to be deleted in all of the patients with the 22q13 deletion syndrome. Over 50 patients were tested to determine if the region of the NT microdeletion is deleted.
3. In several deletion syndromes a phenotype/genotype correlation can be made: the larger the deletion, the more severe the phenotype. There has not been a large enough sample size of the 22q13 deletion syndrome molecularly characterized to make such a correlation. Samples from approximately 50 unpublished patients with a deletion of 22q13 along with several previously published patients were obtained to determine the size of the deletion and whether there is a correlation with the severity of the phenotype. Once the sizes of the deletions are determined



clinical features associated with the 22q13 deletion syndrome can be mapped.

Mapping the clinical features of the syndrome would allow for examination of the region for potential candidate genes for these other clinical features.





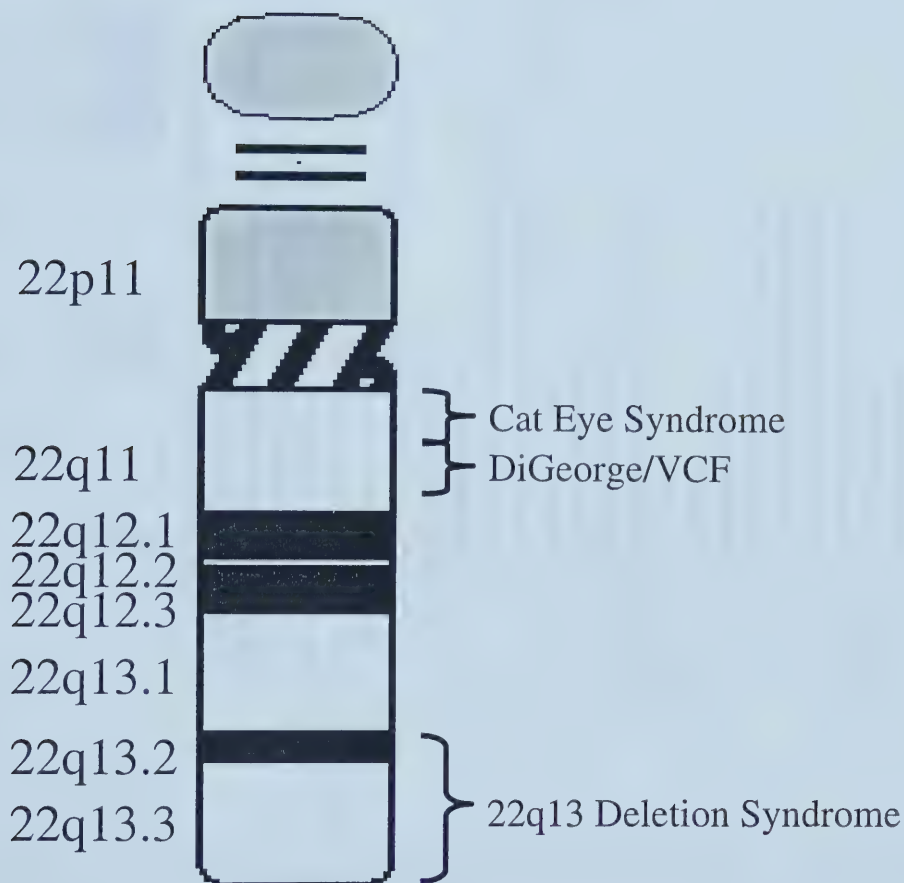


Figure 1-1: Ideogram of Chromosome 22. Band names are identified on the left side of the ideogram. Regions of some of the better characterized chromosome abnormalities associated with chromosome 22 are indicated on the right side.



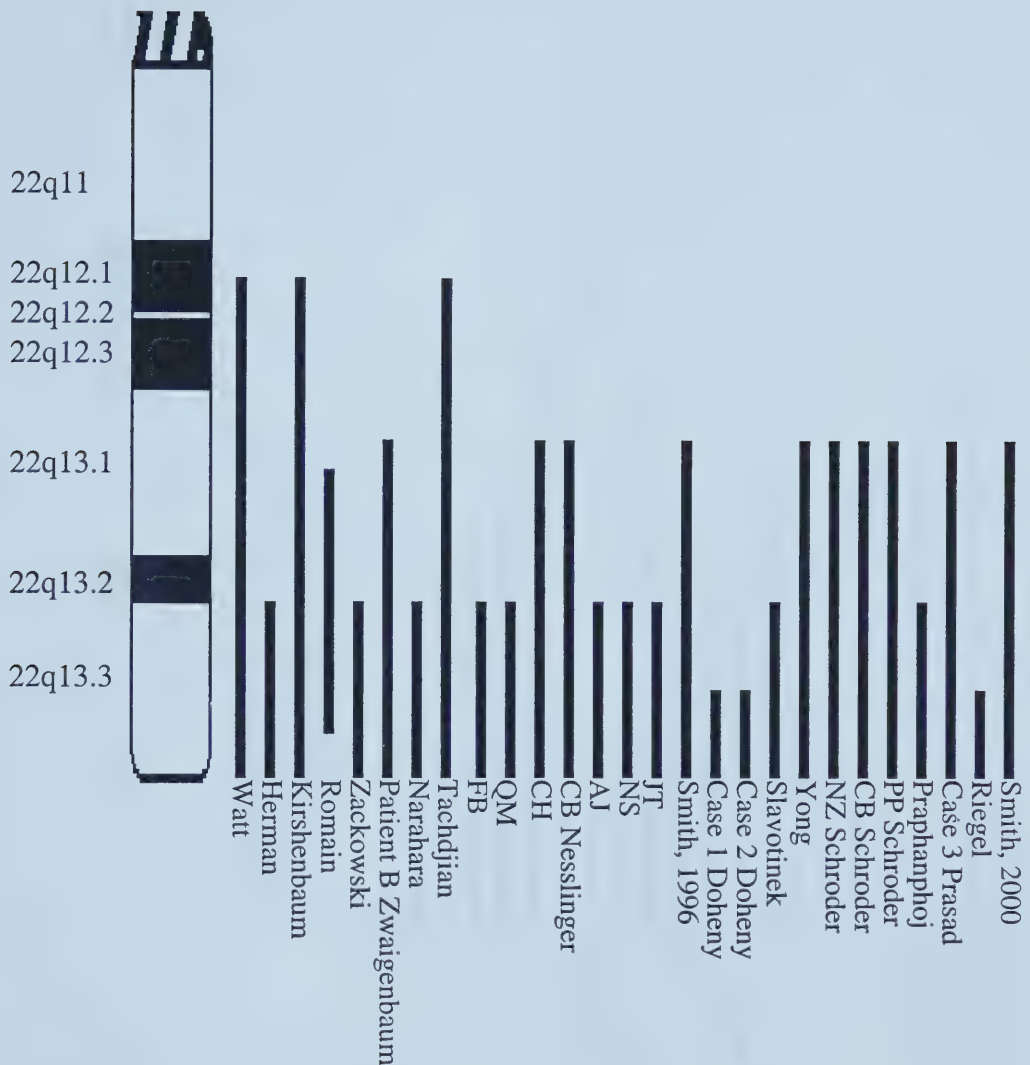


Figure 1-2: Diagram of cytogenetic information of published patient deletions. Left side shows an ideogram of the long arm of chromosome 22 with bands identified. Each vertical line represents the size of the deletion in a published patient based on cytogenetic information. Patients are as follows: Watt, et al, 1985; Herman, et al, 1988; Kirshenbaum, et al, 1988; Romain, et al, 1990; Zackowski, et al, 1990; Patient B from Zwaigenbaum et al, 1990; Narahara, et al, 1992; Tachdjian, et al, 1992; FB from Phelan, et al, 1992 and Nesslinger, et al, 1994; QM from Zwaigenbaum, et al, 1990 and Nesslinger et al, 1994; CH, CB, AJ, NS and JT from Nesslinger, et al, 1994; Smith, et al, 1996; Case 1 and Case 2 from Doheny, et al, 1997; Slavotinek, et al, 1997; Yong, et al, 1997, NZ, CB and PP from Schroder, et al, 1998; Praphanphoj, et al, 2000; Case 3 from Prasad, et al, 2000; Riegel, et al, 2000; Smith, et al, 2000.



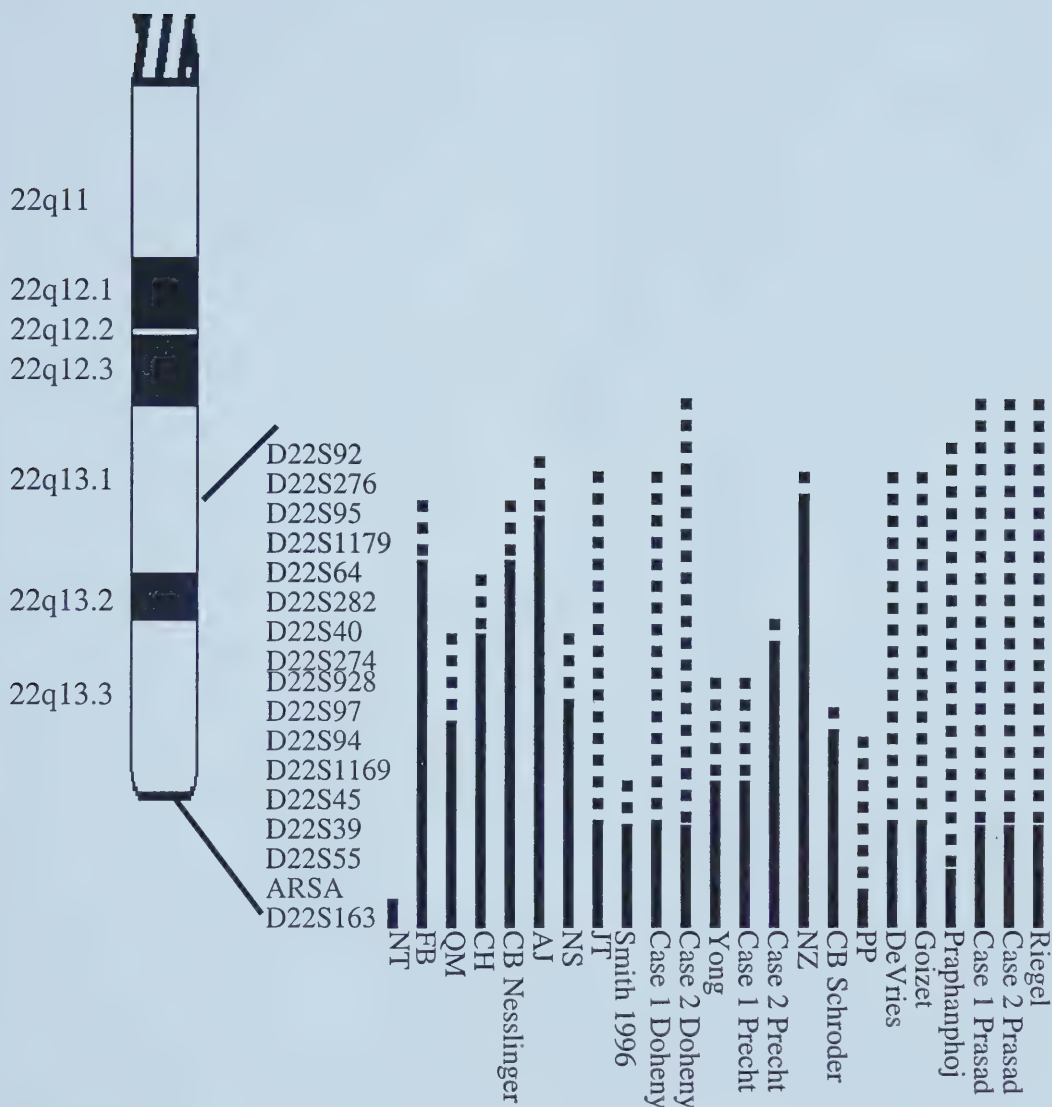


Figure 1-3: Diagram of molecular information of published patient deletions. Markers are identified in order from proximal to distal and approximate location on the cytogenetic map are shown. Each vertical line represents a single patient's deletion. A solid line is the region confirmed to be deleted and a dotted line is an unconfirmed region of deletion. Patients are as follows: NT from Flint, et al, 1995; FB, QM, CH, CB, AJ, NS, and JT from Nesslinger, et al, 1994; Smith, et al, 1996; Case 1 and Case 2 from Doheny, et al, 1997; Yong, et al, 1997; Case 1 and Case 2 from Precht, et al, 1998; NZ, CB and PP from Schroder, et al, 1998; DeVries, et al, 2000; Goizet, et al, 2000; Praphanphoj, et al, 2000; Case 1 and Case 2 from Prasad, et al, 2000; Riegel, et al, 2000.



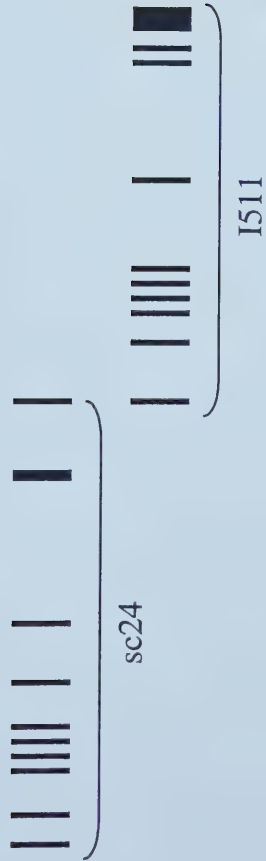




Figure 1-4: Diagram of “*ALPR*” as identified by A. Wong (1998). Each exon is identified by a vertical bar. Solid black bars are confirmed exons, whereas dotted black bars are predicted exons. Exons are drawn in relative size to one another. The names given to each region were assigned by A. Wong. Each grouping of exons represents a cDNA obtained or a predicted group of exons. Starting exon is a single predicted exon and was hypothesized by A. Wong to contain the start methionine codon of this gene. sc24 is a RT-PCR product obtained from spleen by A. Wong that contains most of the 5’ end of the gene. Exon 55337 is an exon that was obtained as part of two cDNAs, sc24 and I511. I511 is a cDNA obtained by A. Wong from a fetal brain library. The genscanex1 is a large exon that was originally predicted by the Genscan exon prediction program. FLS and FL2 are alternative 3’ ends that were identified from ESTs. The Last Exon is an exon predicted by A. Wong to be one of the 3’ ends of the *ALPR* gene (modified from Wong, 1998).

..... Starting exon

Exon 55337



genscanex1

FLS

FL2

≡≡≡ Last Exon







**Table 1-1: Summary of cytogenetic and clinical features of published patients with a deletion of 22q13**

[illegible]



## **Chapter 2: Materials and Methods**

### **Selection of Patients:**

All of the patients in this study were selected based on having a cytogenetic abnormality of the terminal region of chromosome 22q. Nine of the patients in this study had been previously published (1/NT and 10/LM were published in Flint, et al, 1995; 15/CH, 17/AJ, 18/JT, 22/QM, 35/CB, 39/FB and 49/NS were published in Nesslinger, et al, 1994; patient 39/FB was also previously published in Phelan, et al, 1988 and Phelan, et al, 1992; patient 22/QM was also previously published in Zwaigenbaum, et al, 1990). Blood samples or cell lines were obtained on all of the patients and as many of the parents as possible. If the patient had an unusual chromosomal abnormality then some of the blood was sent to produce a cell line. DNA was extracted and FISH cells prepared as described below.

Patients 2, 6, 9, 12, 14, 16, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 31, 32, 34, 35, 38, 39, 40, 41, 42, 44, 46, 47, 48, 49, 50, 52, 53, 54, 55 and JaA were all examined by the same group of clinicians and published in Phelan, et al, 2001. All of these patients and their families are part of the parent support group for the 22q13 deletion syndrome. At the two meetings of this support group many of these patients were evaluated by clinicians and the blood samples were obtained at this time. Developmental assessment of these patients was carried out using the Scales of Independent Behavior (SIB) and Developmental Profile II (DP).

The SIB is a parental interview based approach to assessing the interaction of the child with their social and physical environments. This assessment is broken down into





several sections. The SIB-BI is the broad independence score and is similar to an IQ score (NA indicates that the patient's score was too low to calculate). This score is also presented in an age equivalent in months (what functional age is the patient at) to allow for a more accurate evaluation of the patient's intellectual abilities. The SIB-SS component measures the amount of support the patient requires. This component is broken down as follows: a score of 1-24 means that the patient requires full 24 hour one on one support, 25-39 means that the patient needs one on one support during the day, 40-54 indicates that the patient needs frequent support, 55-69 the patient requires limited support, 70-84 the patient requires intermittent support; and a score of 85-100 means the patient requires infrequent or no support. The other components of the SIB test are all reported in age equivalents and include: SIB-MS which is an evaluation of the motor skills; SIB-SC which measures the social and communication skills; the SIB-PL which measures the personal living skills; and the SIB-CL which is an evaluation of the community living skills.

The DP test is also an age equivalent based assessment of the patients' developmental skills. Patients are evaluated in this test through direct observation and assessment, as well as parental interviews. DP P is the developmental profile for the physical skills of the patient. Self-help skills are assessed in the DP SH component of this evaluation. The patient's social skills are evaluated in the DP S component. The academic skills of the patient are assessed in the DP A component. The DP C component evaluates the communication skills of the patient. The DP IQ score is an estimated intelligence score based on the previous components.



The remaining patients in the study were obtained through contact with the parents, doctors or other clinicians. The clinical features in these patients were obtained through examination of the medical records, and the parent's comments. Thus, the clinical features of these patients were not as well characterized.

### **Isolation of Nucleic Acids:**

#### **DNA Extraction from Whole Blood (Modification of Gustincich, et al, 1991):**

Two volumes of Blood Lysis Buffer (8% DTAB (Sigma), 1.5 M NaCl, 100 mM Tris-HCl pH8.6, 50 mM EDTA) was added to the blood blood sample and then incubated at 68°C for 5 minutes and stored at -70°C. The Blood/DTAB mixture was thawed at 68°C and heated for an additional 5 minutes. The heated Blood/DTAB mixture was added to an equal volume of chloroform in a 50mL Oak Ridge tube and mixed by inversion. The mixture was centrifuged at 10 000 rpm in an HB-4 rotor in a Sorvall® RC-5B Refrigerated Superspeed Centrifuge (DuPont Instruments) for 10 minutes at room temperature. The supernatant was transferred to a fresh tube and an equal volume of water and 0.13 volumes of 5% CTAB solution (5% CTAB (Sigma) and 0.4 M NaCl) was added and mixed by inversion. The mixture was centrifuged at 10 000 rpm in the HB-4 rotor in the Sorvall for 10 minutes at room temperature. The supernatant was removed and the remaining pellet resuspended in 5 mL of 1.2 M NaCl overnight at room temperature. 12.5 mL of cold 95% ethanol was added and mixed by inversion. If the DNA was visible it was removed and transferred to a 1.5 mL microfuge tube, where it was pelleted at 14 000 rpm in a benchtop microfuge for 10 minutes followed by a wash with 70% ethanol, and then resuspended in an appropriate volume of sterile TE buffer (10



mM Tris-HCl pH 8.0, 1 mM EDTA) overnight at room temperature. If no DNA was visible the Oak Ridge tube was centrifuged at 10 000 rpm in the HB-4 rotor in the Sorvall for 10 minutes at 4°C, and the pellet was resuspended in an appropriate volume (50 to 500  $\mu$ L) of sterile TE buffer (10 mM Tris-HCl pH8.0, 1 mM EDTA) overnight at room temperature. The concentration of DNA was estimated by gel electrophoresis on an agarose gel.

#### Isolation of Genomic DNA from Cell Line:

Cells were cultured in 10 mL of RPMI 1640 culture media (Roswell Park Memorial Institute 1640 media from GIBCO BRL<sup>®</sup>) supplemented with 10% fetal bovine serum (GIBCO BRL<sup>®</sup>) and 1% l-glutamine (GIBCO BRL<sup>®</sup>) in a 25 cm<sup>2</sup> cell culture flask (Corning), by P. Brinkman-Mills. Cells were grown at 37°C and 5% carbon dioxide. Five flasks of cells were pooled together and centrifuged at 1000 rpm for 10 minutes. All but 5 mL of the media was removed, and the pellet was resuspended in the remaining volume. 10 mL of Blood Lysis Buffer (8% DTAB (Sigma), 1.5 M NaCl, 100 mM Tris-HCl pH8.6, 50 mM EDTA) was added and the mixture was heated to 68°C for 20 minutes and stored at -70°C. The cell/DTAB mixture was thawed at 68°C and heated for an additional 5 minutes. The heated cell/DTAB mixture was added to an equal volume of chloroform in a 50mL Oak Ridge tube and mixed by inversion. The mixture was centrifuged at 10 000 rpm in an HB-4 rotor in a Sorvall<sup>®</sup> RC-5B Refrigerated Superspeed Centrifuge (DuPont Instruments) for 10 minutes at room temperature. The aqueous layer was carefully removed and transferred to a fresh 50mL Oak Ridge tube where an equal volume of chloroform was added and mixed by inversion. The chloroform mixture was



centrifuged at 10 000 rpm in the HB-4 rotor in the Sorvall centrifuge for 10 minutes at room temperature. The remainder of the protocol was identical to that for isolation of DNA from whole blood (above).

#### Isolation of Plasmid DNA from Bacterial Cells (Modification of Invitrogen®

#### Corporations's Bacterial Alkaline Lysis Miniprep protocol):

Single colonies were selected and a 5 mL culture was set up (5 mL of LB broth and 10  $\mu$ g of ampicillin) and incubated at 37 °C overnight. 1.5 mL of the culture was transferred to 1.5 mL microfuge tubes and centrifuged at 14 000 rpm for 1 minute. The LB supernatant was poured off leaving a few drops remaining. This was repeated with another 1.5 mL of the culture. The resulting pellet was resuspended in the remaining LB by vortexing. 300  $\mu$ L of freshly made TENS buffer was added to the bacterial cells (1X TE, 1% SDS, 0.2M NaOH) and vortexed well to mix. 300  $\mu$ L of 3 M NaOAc was added to the mixture and vortexed. The lysed cells were centrifuged at 14 000 rpm for 5 minutes to remove bacterial debris. The resulting supernatant was transferred to a new 1.5mL microfuge tube and 700  $\mu$ L of phenol:chloroform:isoamyl alcohol (25:24:1) was added and the mixture was vortexed. The phenol:chloroform mixture was centrifuged at 14 000 rpm for 5 minutes and the aqueous phase was removed and transferred to a new tube. 500  $\mu$ L of chloroform:isoamyl alcohol (24:1) was added and the mixture vortexed. The mixture was centrifuged at 14 000 rpm for 5 minutes and the aqueous phase was transferred to a new tube. 900  $\mu$ L of ice cold 95% ethanol was added and this was left at -70°C for at least 20 minutes. The DNA was precipitated by centrifuging at 14 000 rpm for 15 minutes and the 95% ethanol was removed. 200  $\mu$ L of 70% ethanol was added to





wash the DNA pellet and this was centrifuged at 14 000 rpm for approximately 2 minutes. The ethanol was removed and the pellet was allowed to air dry. The pellet was then resuspended in an appropriate volume (20 to 100  $\mu$ L) of TE with 10  $\mu$ g of RNase A added.

#### Isolation of Cosmid DNA for FISH probes:

Cosmids from a frozen chromosome 22 cosmid library were streaked on LB plates containing 400  $\mu$ g of kanamycin antibiotic and incubated at 37°C overnight. A single colony from the plate was selected and grown in 5 mL of LB media containing 100  $\mu$ g of kanamycin for approximately 8 hours at 37°C with shaking. The 5 mL culture was used to inoculate 500 mL of LB media containing 5 mg of kanamycin and grown overnight at 37°C with shaking. The bacterial cells were pelleted using a Sorvall® RC-5B Refrigerated Superspeed Centrifuge (DuPont Instruments) at 5000 rpm for 10 minutes at 4°C in either a GSA or SS-34 rotor. Cosmid DNA was isolated using QIAGEN® Maxi Kit with QIAGEN®-tip 500 and following the protocol provided by supplier with a few modifications. The bacterial pellet was resuspended by vortexing in 10 mL of buffer P1 (50 mM Tris-HCl, pH 8.0, 10 mM EDTA). 10 mL of the lysis buffer P2 (200 mM NaOH and 1% SDS) was added and mixed by inversion followed by an incubation for 5 minutes at room temperature. Then 10 mL of the neutralization buffer P3 (3 M KOAc) was added and mixed by inversion followed by an incubation on ice for at least ½ hour. This mixture was centrifuged at 9500 rpm in the Sorvall Centrifuge for ½ hour at 4°C. The supernatant was passed through a layer of gauze before entering the equilibrated (with 10 mL of buffer QBT; 750 mM NaCl, 50 mM MOPS pH 7.0, 15% ethanol, 0.15% Triton



X-100) QIAGEN®-tip 500. The QIAGEN®-tip 500 was washed twice with 30 mL of buffer QC (1 M NaCl, 50 mM MOPS pH 7.0, 15% ethanol). The cosmid DNA was eluted from the column using 15 mL of buffer QF (1.25 M NaCl, 50 mM Tris-HCl pH 8.5, 15% ethanol). The DNA was precipitated by adding 0.7 volumes of room temperature isopropanol and mixing by inversion. The DNA was pelleted by centrifuging at 9500 rpm for ½ hour at 4°C in the Sorvall Centrifuge in the SS-34 rotor. The ethanol was removed and the DNA pellet was washed in 5 mL of 70% ethanol and centrifuged in the Sorvall Centrifuge at 9500 rpm for 10 minutes at 4°C. The ethanol was removed and the pellet allowed to air dry for approximately 10 minutes. The pellet of cosmid DNA was then resuspended in 500 µL of TE (10mM Tris-HCl pH 8.0, 1mM EDTA) overnight at room temperature. The concentration of cosmid DNA was estimated by agarose gel electrophoresis.

QIAGEN®-tip 100s were also used when a lower yield of cosmid DNA was needed. For these cosmids, the overnight culture was grown in 100 mL of LB with 1 mg of kanomycin added. The protocol followed was the same with the exception of the volumes used. 4 mL of buffers P1, P2, P3, and QBT were used with these smaller columns. 10 mL was used in each of the washes of the column with buffer QC and the cosmid DNA was eluted in 5 mL of buffer QF. The pellet was washed in 2 mL of 70% ethanol and then resuspended in 100 µL of TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA) overnight at room temperature.

#### Isolation of total RNA from tissue samples:

The tissue sample was weighed to obtain the weight (target mass usually 0.5 g to



1 g) so that an appropriate volume of TRIzol<sup>®</sup> (GIBCOBRL<sup>®</sup>) could be added in a Corning 50mL tube (1 mL for mass up to 500 mg and 1.6 mL for mass up to 1 g). The tissue was then homogenized until no large pieces remained (following protocol provided by GIBCOBRL<sup>®</sup>). The samples were then transferred to a 2 mL microfuge tube and centrifuged in a benchtop microfuge at 14 000 rpm for 10 minutes at 4°C. The supernatant was transferred to a new tube and 320  $\mu$ L of chloroform was added and mixed thoroughly followed by a 2 minute incubation at room temperature. This was centrifuged at 11 000 rpm for 15 minutes at 4°C and the aqueous layer was transferred to a fresh microfuge tube. 960  $\mu$ L of ice cold isopropanol was added and mixed by inversion followed by a room temperature incubation for 10 minutes. The RNA was pelleted by centrifuging at 11 000 rpm for 15 minutes at 4°C, the isopropanol was removed and the liquid allowed to drain. The RNA pellet was resuspended in 600  $\mu$ L of RNA Isolation Buffer (4 M guanidine thiocyanate, 0.05% N-lauroyl-sarcosine, 25 mM sodium citrate), and the RNA was again precipitated by adding 1.2 mL of ice cold ethanol followed by a room temperature incubation for 10 minutes. The RNA was pelleted by centrifuging at 11 000 rpm for 15 minutes at 4°C. The isopropanol was removed and the RNA pellet was washed in 1 mL of ice cold 75% ethanol, followed by a centrifugation at 11 000 rpm for 10 minutes at 4°C. The ethanol was removed and the RNA pellet was allowed to air dry for approximately 5 minutes. The RNA pellet was then resuspended in an appropriate volume of DEPC treated water and was stored at -70°C until needed.

#### Isolation of mRNA from total RNA:

mRNA was isolated from total RNA using the Promega PolyATtract<sup>®</sup> System IV



following the manufacturer's protocol. Approximately 200  $\mu$ g of total RNA was heated to 65°C and 3  $\mu$ L of the Biotinylated-Oligo(dT) probe was added with 13  $\mu$ L of 20X SSC and the mixture was allowed to cool to room temperature. The Streptavidin-Paramagnetic Particles (SA-PMP) were resuspended in the storage buffer and then were allowed to pellet on the provided magnetic stand. The supernatant was removed and the pellet of SA-PMPs was washed 3 times in 0.5X SSC (0.3 mL per wash). The washed SA-PMPs were then re-suspended in 0.1 mL of 0.5X SSC and the RNA-probe mixture was added. The mixture was incubated at room temperature for 10 minutes, mixing every 1 to 2 minutes. The SA-PMPs were pelleted using the magnetic stand and the supernatant was carefully removed. The pellet was then washed 4 times in 0.1X SSC (0.3 mL per wash). The mRNA was eluted from the SA-PMPs into 0.1 mL of RNase-Free water, and the supernatant was transferred to a fresh microfuge tube. The elution was repeated with 0.15 mL of RNase-Free water and the supernatant was pooled with the first elution step. The mRNA was then concentrated using a speed-vac and the final concentration and purity of the mRNA was determined using a spectrometer.

### **Fluorescent *In Situ* Hybridization (FISH):**

Preparation of Metaphase Cells for use in FISH (Modified from Arakai, and Sparkes, 1963):

6 to 10 drops of whole blood was added to 10 mL of RPMI 1640 culture media (Roswell Park Memorial Institute 1640 media from GIBCOBRL®) supplemented with 10% fetal bovine serum (GIBCOBRL®) and 1% l-glutamine (GIBCOBRL®) in a 25 cm<sup>2</sup> cell culture flask (Corning). 0.2 mL of phytohemagglutinin M-form (PHA from GIBCOBRL®)





was added and the cell cultures were incubated at 37°C in a 5% carbon dioxide incubator for 72 to 96 hours. When cells from either blood or cells lines were ready 2 µg of KARYOMAX® COLCEMID® (GIBCOBRL®) was added and the cultures were incubated at 37°C for 25 minutes to arrest cells in metaphase. Cultures were transferred to a centrifuge tube and centrifuged at 900 rpm for 10 minutes at room temperature in a Jouan B 3.11 centrifuge. The supernatant was removed and the cell pellet was resuspended in 5 mL of pre-warmed, hypotonic (0.075 M) KCl and then incubated at 37°C for 18 minutes. 5 to 7 drops of fixative (3 methanol:1 glacial acetic acid) were added to the cells and was gently mixed by inversion. The cells were centrifuged at 900 rpm for 10 minutes at room temperature and the supernatant was removed. 5 mL of fixative was added and the pellet carefully resuspended by inversion. Cell preparations were kept at 4°C for at least 24 hours (up to a week). Cells were pelleted by centrifugation at 900 rpm for 10 minutes at room temperature and the supernatant removed. The pellet was resuspended in 5 mL of fixative and this wash with fixative was repeated two more times, with the last resuspension of the cell pellet in an appropriate volume (1 to 5 mL) of fixative. Cells were stored at -20°C until needed.

#### Labeling FISH Probes:

BAC, PAC or cosmid DNA was labeled using GIBCOBRL® BioNick™ Labeling system following the protocol provided by the manufacturer. Approximately 2-3 µg DNA was used in the probe labeling reactions with 1x dNTP mix (containing biotin-14-dATP) and 1x enzyme mix (some labeling reactions contained an additional 5 to 10 units of DNA polymerase I, from GIBCOBRL®, to improve the incorporation of the biotin labeled



nucleotide). Incubation at 16 °C was carried out for one and a half to two hours. The reactions were stopped by freezing the samples at -20°C. Concentration of the probe and the average size of the fragments was determined by agarose gel electrophoresis on a 2% agarose gel. Ideal fragment size was between 100 and 500 base pairs. If the average size of the fragments were too large, more 10x enzyme mix was added and the reactions incubated at 16°C for an additional 15 minutes to half and hour. Labeled reactions were stored at -20°C until needed.

The efficiency of the labeling of some of the probes was assayed using GIBCOBRL® BLUGENE® Nonradioactive Nucleic Acid Detection System. 2.5 µL of the labeling reaction was dotted on HyBond-N membranes (Amersham Pharmacia Biotech UK Limited) with several dilutions of the labeling reaction in water. After baking the membrane at 65°C for one hour, the membrane was rehydrated in Buffer 1 (100 mM Tris pH 7.5 and 150 mM NaCl) for one minute. A blocking step was carried out by incubating the membrane in Buffer 2 (100 mM Tris pH7.5, 150 mM NaCl and 3% BSA (Fraction V, Sigma) at 65°C for one hour. The membrane was then incubated in 0.6 mg of SA-AP (streptavidin-alkaline phosphatase) conjugate diluted in Buffer 1 at room temperature for 10 minutes followed by two washes of 10 minutes in Buffer 1 at room temperature. The membrane was then washed for 10 minutes at room temperature in Buffer 3 (100 mM Tris pH9.5, 100 mM NaCl and 50 mM MgCl). The signal was visualized using 4.9 mg of NBT (Nitroblue tetrazolium) and 2.5 mg BCIP (5-bromo-4-chloro-3-indolylphosphate) in 100 mM Tris pH 9.5, 100 mM NaCl and 50 mM MgCl and incubated at room temperature in the dark until the spots were visible (30 minutes to overnight). Reactions



were terminated by washing membrane in TE (10mM Tris-HCl pH 8.0, 1mM EDTA). If the labeling of the DNA was efficient, then dark spots were visible for most of the dilutions of the labeling reactions spotted on the membrane.

Hybridization and Detection of FISH Probes (Modification of McTaggart, et al, 1998):

Metaphase cell preparations were dropped onto clean glass microscope slides (Fisher Scientific Co.) and dried slowly (either over a 78°C water bath or on moist paper towels at room temperature) to allow for the spreading of the chromosomes. The slides were then treated with RNase A, 150  $\mu$ L (0.5  $\mu$ g/ $\mu$ L) for each 22x22mm area to be used. These were incubated at 37°C for ½ hour in a humid environment (either in a container with a sponge in an air incubator, or in a MJ Research Slide Tower on the MJ Research Peltier Thermal Cycler-200 base, or in the In-Slide-Out incubator from Boekel). After treatment with RNase A, the slides were washed three times in 2X SSC at room temperature for 2 minutes each time. The slides were then dehydrated through a series of ethanol washes (70%, 95% and 100%) for 5 minutes in ice cold ethanol. Following dehydration, the slides were allowed to air dry before the denaturation step. To denature the chromosomes on the slides, the slides were incubated for two minutes at 72°C in 70% formamide (Sigma) and 2X SSC. Immediately following the denaturation the slides were dehydrated as above and allowed to air dry before the probe was added.

The probes for FISH were prepared by adding approximately 500 ng of the biotin labeled DNA to 8  $\mu$ g of Human Cot-1 DNA (GIBCOBRL®). This was precipitated by adding 0.1 volumes of 3 M NaOAc and 2.5 volumes of 95% ethanol, and mixing well. This was stored at least one hour (preferably overnight) at -70°C, and was then



centrifuged in a benchtop microfuge at 14 000 rpm for 20 minutes at room temperature. The resulting pellet was washed in 200  $\mu$ L of 70% ethanol and put at -70°C for at least a ½ hour. The probes were then centrifuged, the ethanol removed and the pellet was allowed to air dry. Once dry the pellet was resuspended in 20  $\mu$ L of HybrisolVII (50% formamide, 2X SSC from ONCOR® Appligene) at 42°C for 15 minutes, vortexing frequently. Once the probes were resuspended, they were denatured at 75 to 80°C for 10 minutes, followed by reannealing at 37°C for at least one hour. The probes were added to glass coverslips and the slides were then brought down on top of the coverslips. The edges of the coverslip were sealed with rubber cement and the slides were hybridized for 16 to 20 hours at 37°C in a humid environment.

The next day, the rubber cement and the coverslips were carefully removed from the slides and the slides were immediately placed into a 50% formamide/2X SSC wash at 45°C for 5 minutes. This wash was repeated three more times followed by three washes in 2X SSC at 45°C for 5 minutes each. 40  $\mu$ L of blocking reagent (4X SSC, 0.1% Tween-20, 3% BSA (Fraction V, Sigma®)) was added to each 22x22mm area of the slide and the slides were incubated at 37°C for at least 15 minutes covered by a plastic coverslip. The coverslips were removed and 40  $\mu$ L of FITC-avidin (ONCOR® Appligene, or Vector Laboratories 10  $\mu$ g/mL) was added to the slides. The slides were covered with plastic coverslips and incubated at 37°C for at least a ½ hour. The slides were then washed three or four times in 4X SSC, 0.1% or 0.2% Tween-20 at 45°C for 5 minutes each. After the washes the blocking step was repeated, followed by an incubation with 40





$\mu$ L of anti-avidin antibody (ONCOR<sup>®</sup> Appligene, or Vector Laboratories biotinylated anti-avidin 6.25  $\mu$ g/mL) at 37°C for at least a ½ hour. Again three or four washes (4X SSC, 0.1% or 0.2% Tween-20) at 45°C for 5 minutes each, were carried out. The blocking step was repeated for a third time followed by another incubation with FITC-avidin at 37°C for at least a ½ hour. This incubation was followed by a final series of washes (4X SSC, 0.1% or 0.2% Tween-20) at 45°C for 5 minutes each. Most of the liquid was removed from the slides and approximately 20 to 25  $\mu$ L of Propidium Iodide in a mounting medium (ONCOR<sup>®</sup> Appligene, or Vector Laboratories) was added to the slides. A coverslip was added and excess moisture was removed from the slides before sealing the edges with nail polish.

The slides were stored in the dark until ready for viewing on a Zeiss Axiophot microscope using a triple band filter to allow for visualization of the red and green colors used in the FISH analysis. Photographic slides were taken of several metaphase preparations for each patient using Kodak EliteChrome Color Slide Film (ASA 100).

### **Southern and Northern Blot Analysis:**

#### **Preparation of Southern Blots:**

DNA samples were prepared as needed and run on an appropriate concentration of agarose gel. For plasmid, or PCR DNA samples, the gel was washed in 0.4 M NaOH for ½ hour and the transfer was set up using a modification of Southern's (1975) protocol, using 0.4 M NaOH as the transfer buffer. The membranes used were GeneScreen Plus<sup>®</sup> from NEN<sup>™</sup> Life Science Products. After the overnight transfer the membrane was



washed in 0.4 M NaOH.

For genomic DNA the gel was first denatured in 0.5 M NaOH and 1.5 M NaCl two times for 20 minutes each, this was followed by a neutralization step of 0.5 M Tris and 3 M sodium chloride for 20 minutes repeated once (modified from Sambrook, et al, 1989). Again the transfer was set up using a modification of Southern's (1975) protocol using 10X SSC as transfer buffer and GeneScreen Plus® membranes. After transferring overnight the membrane was removed and washed in 0.4 M NaOH for 5 minutes followed by a wash of 0.2 M Tris-Cl and 2X SSC (pH 7.5) for at least 15 minutes. After these washes, the membrane was cross-linked using ultra-violet light exposure for approximately 20 seconds. The Southern Blots were then probed with double or single stranded DNA probes.

#### Preparation of Northern Blots:

Approximately 40 µg of total RNA was added to 1X MOPS, 50% formamide and 18% formaldehyde, and incubated at 60°C for 15 minutes before loading on a 1.2% agarose gel (modified from Sambrook and Russell, 2001). The agarose gel was prepared in DEPC treated water with 1X MOPS and 5% formaldehyde and pre-run in 1X MOPS for 5 -15 minutes. 1/5th volume of an RNA loading buffer (50% glycerol, 10 mM EDTA, 0.25% bromophenol blue, 0.25% xylene cyanol) was added to the heated RNA samples prior to loading. Gels were run at an appropriate voltage for the size of the gel and for an appropriate amount of time. After the gel was run, the RNA ladders were removed, stained in Ethidium Bromide and photographed with a ruler to allow for sizing of fragments. The remainder of the gel with the RNA samples was rinsed in DEPC treated



Milli-Q water for 45 minutes to remove the formamide and formaldehyde. The gel was then washed in 10X SSC (treated with DEPC) for 45 minutes and membrane transfer was set up using 20X SSC (treated with DEPC) as the transfer buffer and allowed to transfer overnight. The membrane was then removed and baked at 80°C for 2 hours and stored at 4°C until needed.

#### Labeling of Radioactive Probes for Northern and Southern Blots:

Probes were labeled using standard oligo-labeling or either the Ambion® Strip-EZ™ DNA or PCR kits. For oligo-labeling the DNA template was boiled for 10 minutes to denature before adding to 5 µL of OLB/ $\beta$ -mecaptoethanol, 1 µL BSA (10X BSA), and 25 µCi  $\alpha$ [<sup>32</sup>P]dCTP (modified from Sambrook et al, 1989). Approximately 10 units of Klenow DNA polymerase was added and the reaction was incubated at 37°C for at least one hour. For the Ambion® Strip-EZ™ DNA kit for a 25 µL probe reaction the manufacturer's protocol was followed. The template DNA was denatured (boiled for 10 minutes) and was added to 1X Decamer solution (Ambion®), 1X Buffer -dATP/-dCTP (Ambion®), 1X dCTP, 50 µCi of  $\alpha$ [<sup>32</sup>P]dATP, and 1 µL of Exonuclease-free Klenow (Ambion®). The labeling reaction was incubated at 37°C for ½ to 1 hour, the reactions were stopped by adding 50mM EDTA, 15% Glycerol and blue dextran. For the Ambion® Strip-EZ™ PCR kit the manufacturer's protocol was followed. A 20 µL PCR reaction was carried out in 1X PCR buffer (Ambion®), 1X dNTP solution (Ambion®), 20 µCi of  $\alpha$ [<sup>32</sup>P]dATP, 10 pmol of antisense primer, an appropriate concentration of DNA template and Taq:pfu (ratio 25:1, approximate concentrations Taq 5 U/µL and Pfu 2 U/µL) (Microbiology, University of Alberta). PCR was carried out in a Progene thermal cycler



(Techne) denaturing the template at 95°C for 30seconds to 1 minute, an appropriate annealing temperature for the primer and an extension at 72°C for an appropriate length of time (usually 1 minute for 1kb) for 30 to 40 cycles.

Unincorporated nucleotides were removed from the reactions by passing labeling reaction over Sephadex™ G-50 Fine columns. If the probe contained repetitive sequences, then 84  $\mu$ L of 1M NaPO<sub>4</sub> and 500  $\mu$ L of sonicated placental DNA were added and the mixture boiled for 10 minutes and then allowed to pre-anneal at 65°C for at least one hour. If no repetitive sequences were contained in the probe then 2mg of herring sperm DNA was added to the probe and the mixture was boiled for 10 minutes before adding to the blot. Single stranded PCR probes were not boiled, but were added to the blots immediately after removing the unincorporated nucleotides.

#### Southern Blot Hybridization:

The Southern blot was prehybridized in an appropriate volume of a modified Church and Gilbert solution for at least one hour at 65 °C in a hybridization oven (Tyler Research Instruments). The probe was added and the hybridization allowed to proceed at 65 °C for 16 to 24 hours. The next day the blots were washed twice in a low stringency wash (1.5X SSC, and 0.2% SDS) at room temperature for 10 minutes, followed by at least one wash in a high stringency wash solution (0.2X SSC, and 0.2% SDS) at 65 °C for 10 minutes. If the blot still appeared to have a lot of background when checked with the Geiger counter, then the blot was washed in a super high stringency wash (0.1X SSC, and 0.2% SDS) at 65 °C for at least 10 minutes. Blots were sealed in plastic and exposed to X-ray film (FUJI Medical X-Ray Film Super RX or Kodak X-OMAT™ AR Scientific





Imaging Film) for an appropriate amount of time (usually overnight to one week), while stored at -70 °C.

#### Northern Blot Hybridization:

Northern blots were either purchased from Clontech Laboratories Inc. or were made from RNA isolated from tissue samples. The Northern blots were prehybridized for a minimum of one hour in a Northern hybridization solution (50% formamide, 0.02% SDS, 0.4 mg/mL herring sperm DNA, 5X SSPE, 10X Denhardt's) at 42°C in either a hybridization oven (Tyler Research Instruments) or a water bath on a shaker. The probe was added to the blots and hybridization was carried out at 42°C in either the hybridization oven or the water bath for a minimum of 16 hours to a maximum of 40 hours. Blots were washed twice for 10 minutes each, in a low stringency wash solution (2X SSC, and 0.1% SDS) at room temperature. The low stringency washes were followed by a high stringency wash (0.1X SSC, and 0.1% SDS) for 10 minutes with the wash buffer heated to 50°C. After this high stringency wash, the blots were checked with a Geiger counter to estimate the amount of background hybridization. If little background was detected, the blots were sealed in plastic and exposed to film (FUJI Medical X-Ray Film Super RX or Kodak X-OMAT™ AR Scientific Imaging Film) for an appropriate amount of time (overnight to several weeks) while stored at -70°C. If background was detected, the high stringency wash was repeated. If background was still detected, the wash was repeated and the temperature was increased.

UltraHyb™ buffer (Ambion®) was also used for some Northern blot hybridizations following the manufacturer's protocol. Northern blots were prepared as



outlined above, using UltraHyb™ instead of the Northern hybridization buffer. The hybridization was carried out at 42°C and the washes were carried out as described above.

#### Stripping Northern and Southern Blots for Re-use:

Blots probed using the Ambion® Strip-EZ™ labeling kits were stripped following the protocol provided by the manufacturer. The blots were incubated with 1X Probe Degradation Dilution Buffer and 1X DNA Probe Degradation Buffer at 68°C for 10 minutes. This degradation was followed by incubation in 1X Blot Reconstitution Buffer and 0.1% SDS at 68°C for 10 minutes. Blots were then sealed in plastic and stored at 4°C until needed.

#### Polymerase Chain Reaction Analysis of DNA Samples:

##### Microsatellite Polymerase Chain Reaction:

Microsatellite MapPair™ primers were selected and ordered from Research Genetics at a 20 pmol/μL concentration (table 2-1). The forward primer of the pair was end-labeled with either radioactive  $\gamma$ [<sup>32</sup>P]dATP or  $\gamma$ [<sup>33</sup>P]dATP using T4 Polynucleotide Kinase following the manufacturer's protocol (FPLCpure from Pharmacia Biotech). To label 100 pmol of primer approximately 10 units of the T4 Polynucleotide Kinase enzyme (stored in 20 mM potassium phosphate pH 7.0, 25 mM KCl, 10 mM 2-mercaptoethanol, 10<sup>-7</sup> ATP and 50% glycerol) was used in a reaction buffer (100 mM Tris-acetate, 100 mM magnesium acetate and 500 mM potassium acetate), with 25 μCi of  $\gamma$ [<sup>32</sup>P]dATP or  $\gamma$ [<sup>33</sup>P]dATP. The reaction was incubated at 37 °C for at least half an hour in either an air incubator or a Progene thermal cycler (Techne). Several forward primers were ordered with a fluorescent label, either HEX or TET from GIBCOBRL® to allow for detection on



an ABI377 Automated Sequencer (Applied Biosystems).

PCR reactions were carried out on patient samples and parental samples if available using a modification of the protocol supplied by Research Genetics. PCR reaction consisted of 25 mM Tris pH 9.0, 50 mM KCl, 20  $\mu$ g BSA, either 1.5 mM magnesium chloride or 3 mM magnesium chloride, 3 pmol of each primer, 10 mM dNTPs mix (2.5 mM of each nucleotide), 4% DMSO, 10 mM ammonium sulphate, approximately 0.01  $\mu$ g of template DNA, and 0.3  $\mu$ L of Taq:pfu (25:1 ratio, approximate concentrations Taq 5 U/ $\mu$ L and Pfu 2 U/ $\mu$ L)(Microbiology, University of Alberta) per 10  $\mu$ L reaction. The reactions were carried out under the following conditions: denature 94°C for 5 minutes followed by 23 to 25 cycles of denature 94°C for 30 seconds, anneal 55 to 60.5°C for 20 to 30 seconds, depending on the optimized annealing temperature for the primer pair and extend 72°C for 20 to 30 seconds, depending on size of the PCR products (usually 1 minute for 1kb). PCR reactions were carried out in either a Perkin Elmer 7600 (radioactive) or a MJ Research Peltier Thermal Cycler-200 (fluorescent) thermocycler.

For radioactively labeled primers 3  $\mu$ L of STOP dye (98% formamide, 10 mM EDTA, and approximately 1 mg/mL of each of bromophenol blue and xylene cynol dyes) was added shortly after PCR cycles finished and reactions were stored at 4°C until loading. Samples were denatured at 94°C for 5 minutes and then kept on ice before loading. The PCR products were loaded on an 8% denaturing polyacrylamide gel (19:1 acrylamide:bis-acrylamide; 5.32 g acrylamide, 0.28 g bis-acrylamide, 32 g urea, 3.5 mL 20X glycerol tolerant buffer, and water to 70 mL; polymerized by adding 1.2 mL 10%



ammonium persulfate and 17  $\mu\text{L}$  of N,N,N',N'-tetramethylethylenediamine), 0.4mm thick on a Model S2 sequencing apparatus from GIBCOBRL<sup>®</sup> (size 30 x 40 cm gel). Gels were run at 60 Watts in 1X glycerol tolerant buffer (90 mM Tris, 30 mM Taurine, 0.5 mM EDTA) for 3 to 5 hours depending on the size of the PCR product and dried for 2½ hours under vacuum. Gels were then exposed to X-ray film (FUJI Medical X-Ray Film Super RX) for an appropriate amount of time (overnight to two weeks) while stored at -70°C. Often several different exposures were taken to obtain a good image of the different samples.

For the fluorescently labeled primers, 1  $\mu\text{L}$  of a 10  $\mu\text{L}$  reaction was mixed with 1  $\mu\text{L}$  of a PCR reaction with another primer pair. The two reactions mixed together had different fluorescent labels and different sized PCR products. An equal amount of formamide loading dye was added to the mixture of reactions (Amersham Pharmacia). 1 $\mu\text{L}$  of this mixture was loaded onto a 6% polyacrylamide gel on the ABI™ Prism™ 377 Automated Sequencer (Applied Biosystems). The gels used were 36cm in length and were run at 2400 scans/hour for 2 ½ hours. A digital image of the gel was captured for future analysis.

#### Reverse Transcription from RNA:

Approximately 1  $\mu\text{g}$  of RNA was treated with 1  $\mu\text{L}$  of DNaseI in 1X DNaseI buffer and 1  $\mu\text{L}$  of RNaseOut (GIBCOBRL<sup>®</sup>) and incubated at room temperature for 15 minutes followed by the addition of 25 mM of EDTA and an incubation at 65°C for 10 minutes. GIBCOBRL<sup>®</sup>'s THERMOSCRIPT™ RT-PCR System was then used to create cDNAs from the RNA template following the manufacturer's protocol. 1X cDNA





synthesis buffer, 5 mM DTT, 1 mM dNTPs, 2.5  $\mu$ M primer (oligo-dT, random hexamers or gene specific) and 12 Units of THERMOSCRIPT™ was added on ice to the DNaseI treated RNA. The reaction was then incubated in a thermal cycler (MJ PTC 100™ Programable Thermal Controller) for the following temperatures and times: 42°C for ½ hour, 50°C for 10 minutes, 53°C for 10 minutes, 55°C for 10 minutes, 57°C for 10 minutes, 60°C for 10 minutes. The reaction was terminated by incubating at 85°C for 5 minutes. The template RNA was removed by adding 2 Units of RNaseH and incubating at 37°C for 20 minutes. The RT products were stored at -20°C until needed.

#### Polymerase Chain Reaction to determine genomic structure and transcripts:

Primers were designed throughout the gene, including several sets of nested primers for RACE PCR reactions. Melting temperatures of the primers were determined using the %GC formula (Appendix A). PCR was carried out with an appropriate volume of template DNA, either genomic, cosmid, bacterial colonies, RT reactions or Clontech Marathon-Ready cDNA for Random Amplification of cDNA Ends (RACE). The PCR reaction consisted of 25 mM Tris pH 9.0, 50 mM KCl, 20  $\mu$ g BSA, either 1.5 mM magnesium chloride or 3 mM magnesium chloride, 5 pmol of each primer, 10 mM dNTPs mix (2.5 mM of each nucleotide), and Taq:pfu (25:1 ratio, approximate concentrations Taq 5U/ $\mu$ L and Pfu 2U/ $\mu$ L)(Microbiology, University of Alberta). Reactions were carried out in either a 25 or 50  $\mu$ L volume. Generally the reactions were carried out under the following conditions: denature 94°C for 5-10 minutes followed by 25 to 50 cycles of denature 94°C for 30 seconds to 1 minute, anneal about 5 to 10°C below the calculated melting temperature for the primer for 30 seconds, and extend 72°C



for about 1 minute for 1kb. PCR reactions were carried out in either a MJ PTC 100™ Programable Thermal Controller or a MJ Research Peltier Thermal Cycler-200 thermal cycler.

#### Extraction of PCR product from agarose gel:

PCR products were run on an appropriate concentration of agarose in TAE buffer. The band wanted was cut from the gel and then the GeneClean II® kit (Bio101 Inc.) was used to remove agarose following the manufacturer's protocol. The weight of the gel was determined, and three volumes of NaI was added and the mixture heated at 45-55°C until the agarose dissolved. An appropriate volume of glassmilk was added according to the manufacturer's protocol. The glassmilk mixture was incubated on ice for at least 5 minutes. The glassmilk with the bound DNA was pelleted by centrifuging for 5 seconds at 14 000 rpm. The glassmilk pellet was washed three times in 700 µL of New Wash solution. The DNA was eluted from the glassmilk in a volume of water equal to the volume of glassmilk used, at 45-55°C.

#### DNA Manipulations:

##### Ligation of DNA into plasmids:

pGEM®-T and pGEM®-T Easy (Promega) plasmids were used for ligation of PCR products. The insert DNA was ligated to the vector (1:4 ratio) in 1x ligation buffer (GIBCOBRL®) and T4 ligase (GIBCOBRL®) according to GIBCOBRL®'s protocol. Reactions were incubated at 4°C for at least 24 hours.

##### Transformation of *E. coli* cells with plasmid DNA:

Competent cells, XL1-Blue, was added to the ligation reaction and incubated on



ice for 20 minutes (Sambrook, et al, 1989). The bacteria were then heat shocked at 42°C for 90 seconds followed by another incubation on ice for 5 minutes. LB broth was added to the cells and the mixture was incubated at 37°C 45 minutes. Transformed cells were plated on LB plates with ampicillin, X-gal and IPTG and incubated at 37°C overnight. The presence of an insert in the vector was determined using colony PCR, using the PCR conditions listed above.

#### Sequencing DNA:

DNA was sequenced using either fluorescent dyes or radioactivity. For radioactive sequencing, the Thermo Sequenase radiolabelled terminator cycle sequencing kit was used (Amersham Life Science) following the manufacturer's protocol. The sequencing reaction was set up using approximately 0.25  $\mu$ g of template DNA, 2  $\mu$ L of reaction buffer (260 mM Tris-HCl pH 9.0, 65 mM  $\text{MgCl}_2$ ), 5 pmol of primer and 8 U of Thermo Sequenase in a volume of 20  $\mu$ L. This mixture was aliquoted into four reactions of 4.5  $\mu$ L each. To each of the four reactions, 2  $\mu$ L of dGTP termination master mix was added (7.5  $\mu$ M dATP, dCTP, dTTP, and dGTP) along with 0.5  $\mu$ L of  $\alpha$ [ $^{33}\text{P}$ ]ddNTP (G, A, T, or C, 450  $\mu$ Ci/mL) added to the appropriate reaction. The cycle sequencing was carried out in a Techne thermal cycler under the following conditions: 15 to 30 cycles of denature 94°C for 30 seconds, anneal at an appropriate temperature for the primers for 30 seconds, and extend 72°C for about 1 minute for 1kb. The PCR reactions were terminated by adding 4  $\mu$ L of STOP dye (98% formamide, 10 mM EDTA, and approximately 1 mg/mL of each of bromophenol blue and xylene cynol dyes), and stored at -20°C until needed.



Radiolabelled sequencing reactions were denatured at 95°C for 5 minutes before loading samples on an 8% denaturing polyacrylamide gel (19:1 acrylamide:bis-acrylamide; 5.32 g acrylamide, 0.28 g bis-acrylamide, 32 g urea, 3.5 mL 20X glycerol tolerant buffer, and water to 70 mL; polymerized by adding 1.2 mL 10% ammonium persulfate and 17  $\mu$ L of N,N,N',N'-tetramethylethylenediamine), 0.4mm thick on a Model S2 sequencing apparatus from GIBCOBRL® (size 30 x 40 cm gel). The gels were run for 2 to 7 hours at 55 Watts in 1X glycerol tolerant buffer, before drying under a vacuum for 2 ½ hours. Gels were then exposed to X-ray film (FUJI Medical X-Ray Film Super RX) for an appropriate amount of time (overnight to two weeks) while stored at -70°C. Often several different exposures were taken to obtain a good image of the different samples.

Automated sequencing was also used that employed the use of fluorescent labels. The LI-COR IR<sup>2</sup> sequencing was carried out using fluorescently labeled primers. The forward and reverse primers were labeled with different fluorescent dyes so both could be used in the same reaction. Reactions were set up using the ThermoSequenase® fluorescent labeled primer cycle sequencing kit (Amersham). Template DNA was mixed with the 1.5 pmol of forward and reverse fluorescent primer. This mixture was aliquoted into four tubes labeled A, C, G, T and 1  $\mu$ L of the appropriate reagent (A, C, G or T) was added. The reactions were carried out under the following conditions: 92°C for 2 minutes followed by 30 cycles of 92°C for 30 seconds, 50°C for 15 seconds, and 70°C for 15 seconds. 1.5  $\mu$ L of STOP buffer/loading dye was added and the reactions were sent over to Medical Genetics (University of Alberta) for analysis on the LI-COR IR<sup>2</sup>.

Automated sequencing was also carried out using the ABI™ Prism™ 377





Automated Sequencer (Applied Biosystems). This technology is based on fluorescently labeled dye-terminators. The sequencing reactions were carried out in 4  $\mu\text{L}$  of sequencing pre-mix (DYEnamic™ ET Terminator Cycle Sequencing Kit Amersham Pharmacia Biotech Limited) with 2 pmol of primer, 2  $\mu\text{L}$  of template DNA and 0.5  $\mu\text{L}$  of pellet paint. The cycling reactions were carried out under the following conditions: 30 cycles of 95°C for 20 seconds, 50°C for 15 seconds, 60°C for 1 minute. The sequencing reactions were precipitated by adding 1  $\mu\text{L}$  of sodium acetate/EDTA buffer and 40  $\mu\text{L}$  of 95% ethanol and centrifuging at 14 000 rpm for 15 minutes in a benchtop microfuge. The DNA pellets were washed using 200  $\mu\text{L}$  of 70% ethanol and centrifuging briefly. The ethanol was removed and the pellet allowed to dry. The pellet was resuspended in 1.5  $\mu\text{L}$  of loading dye with vigorous vortexing. 1  $\mu\text{L}$  of samples were loaded on a polyacrylamide gel poured for the ABI™ Prism™ 377 Automated Sequencer (Applied Biosystems) and the samples were run for 10 hours and data was extracted for further analysis.



Table 2-1: Microsatellites Used in Analysis of Deletions

Microsatellite Locus	Type of Repeat	Size of PCR Products	Genomic Distance from Centromere
D22S526	Tetra	340bp	33235 kb
D22S1169	Di	118 to 134 bp	32799 kb
D22S1170	Di	201 to 212 bp	31724 kb
D22S1149	Di	111 to 117 bp	30053 kb
D22S532	Tetra	185bp	29566 kb
D22S1141	Di	252 to 282 bp	29161kb
D22S928	Di	150 to 184 bp	28918 kb
D22S274	Di	202 to 214 bp	28712 kb
D22S1159	Di	194 to 224 bp	28193 kb
D22S1171	Di	126 to 166 bp	27841 kb
D22S927	Di	130bp	27509 kb
D22S282	Di	144 to 164 bp	27230 kb
D22S1179	Di	168 to 186 bp	27001 kb
D22S417	Tetra	183bp	26513 kb
D22S1166	Di	112 to 140 bp	26462 kb
D22S270	Di	128 to 148 bp	26462 kb
D22S1178	Di	247 to 263 bp	25812 kb
D22S1157	Di	233 to 250 bp	25661 kb
D22S276	Di	241 to 263 bp	25433 kb
D22S279	Di	249 to 258 bp	24444 kb
D22S284	Di	86 to 102 bp	23738 kb
D22S423	Di	215 to 235 bp	23803 kb
D22S1155	Di	203 to 214 bp	23159 kb
D22S272	Di	132 to 150 bp	22506 kb
D22S1177	Di	182 to 192 bp	20684 kb
D22S283	Di	126 to 152 bp	20200 kb

**Legend:** Microsatellite loci used in the analysis of patients are identified with information regarding the type of repeat (Di, Tri, or Tetra nucleotide repeat), the size of the PCR product expected in base pair and the distance from the centromere the microsatellite is in kilobase pairs.



## Chapter 3: Results

### Characterization of the Novel Gene in NT Deletion Region:

#### Assembly of *SHANK3* Sequence:

The A. Wong (1998) partially cloned the *ALPR* gene using RT-PCR, screening cDNA libraries and identifying ESTs. None of these five cDNAs identified by A. Wong overlapped each other and none contained more than 11 exons of this large gene. The sequences of these partial cDNAs were used at the start of this thesis project to search the web-based software BLAST to identify any new ESTs in the database or new orthologues of this gene in other organisms. In 1999a, Boeckers, et al identified a protein, Proline Rich Synapse Associated Protein 1 (ProSAP1; later renamed Shank2), that was localized to the post-synaptic density of neurons in rat brain. The sequence of *Shank2* was used to confirm that it was the previously characterized gene *Cortactin Binding Protein 1* (*CortBP1*), which showed similarity to *ALPR* (Wong, 1998). Through screens of brain cDNA libraries using *Shank2* as a probe Boeckers et al identified a closely related family member *ProSAP2* (later renamed *SH3 and multiple ankyrin repeat domains 3*, or *Shank3*). BLAST searches of the NCBI non-redundant databases identified *ALPR* as the human orthologue *Shank3*.

In order to assemble the complete cDNA sequence of human *SHANK3* the partial cDNAs obtained by A. Wong (1998) were compared to the orthologous gene from the rat. The human and the rat copies of *Shank3* show high similarity at the nucleic acid level (approximately 90% over most of the gene)(figure 3-1). The alignment of the rat cDNA sequence to the human genomic sequence identified part of an exon that was not



previously identified (exon 1 in figure 3-2). This sequence was found to be the first 162 bases of the rat cDNA and contains the putative starting methionine codon. The starting methionine codon is conserved in the human sequence of this exon. A second part of the rat cDNA not found in either the human cDNA or genomic sequence was numbered as exon 11 in human (figure 3-2). However, a possible gap in the genomic sequence may be present where this exon should be located, raising the possibility that this exon is present in humans, but may be in a region that is difficult to clone or sequence. The 60bp of sequence from the rat cDNA was found to be composed of 97% G/C (58/60), thus making this region difficult to PCR or sequence. Several exons identified by A. Wong (1998) are part of the human cDNAs but were not found in the rat cDNA. Three of these exons are alternative 3' ends of the gene (exons 24, 25, and 27). It is possible that these alternative 3' ends are in the rat, but have not yet been identified since no rat genomic sequence is available. A fourth exon was identified as part of the coding region of *Shank3* that is not found in the rat, and thus this exon may or may not be part of the gene (exon 22) (figure 3-2). Having the sequence of the rat cDNA allowed for identification of an additional two regions of *SHANK3* not previously characterized in human cDNAs, and a more complete sequence of the gene.

The sequence of *SHANK3* was assembled into a single cDNA using the sequence of the partial cDNAs from A. Wong and the rat sequence for exon 11. The assembled sequence of *SHANK3* was used to predict the protein sequence. Alignment of the predicted proteins of the rat and human Shank3 shows more than 90% identity over the entire protein (figure 3-3). Several key binding domains were identified by Boeckers, et





al (1999b) and include: ankyrin repeat domain, PDZ domain, SH3 domain, Homer-binding domain, SH3 binding domain, and a sterile alpha motif (SAM) domain. All of these domains are conserved and in the same order in human SHANK3. There is nearly 100% similarity between the rat and human sequence of these domains (figure 3-4). Using either protein sequence, a BLAST search of the NCBI protein database identified several related proteins. These other proteins are all members of the Shank family of proteins. As previously mentioned Boeckers, et al. (1999a) identified Shank2 (also called ProSAP1 and CortBP1), and Naisbett, et al. (1999) identified Shank1A (also call SSTRIP and Synamon). These Shank family members show high amino acid sequence similarity in the protein interacting domains (60 to 79% similar), whereas other regions are less conserved (20 to 50% similar). The Shank2 protein is lacking the ankyrin repeats found at the N-terminus in the other two proteins, but all of the other protein interacting domains are present and in the same order. Therefore the *ALPR* gene found in the region of the NT deletion, was identified as a member of the Shank family of proteins and renamed *SHANK3*.

#### Characterization of *SHANK3* Transcripts:

RT-PCR analysis was used to join the different regions of *SHANK3* transcripts previously identified by A. Wong (1998) and to verify the exons found in the rat, but not human cDNA. Different combinations of primers were used on a variety of RT products to try and elucidate the structure of this gene. Kidney, muscle, heart, HeLa, CaCo and fetal brain RNA were used as a template in RT reactions, using either random hexamers, oligo-dT or gene specific primers to prime the synthesis reaction (products summarized in



table 3-1). RT-PCR with heart and muscle between exons 1 and 5 produced a band of the correct size and sequence, indicating that all of these exons are present in transcripts (cDNA start to sc24 figure 3-2). Exons 2 through 10 (cDNA sc24) and 12 through 21 (cDNA I511) were obtained through RT-PCR analysis in heart and muscle. These two regions of *SHANK3* were joined together in an RT-PCR product that contains exons 8, 9, 10, 14, and 15 (cDNA sc24 to I511). This indicates that the exons 11, 12 and 13 were spliced out in at least some of the transcripts. RT-PCR products from heart and muscle also joined exon 23 to exons 25 and 27 (cDNA Gen to FL2), indicating that these exons are part of some transcripts. Exon 24, one of the alternative 3' ends, was confirmed as being present in a transcript using RT-PCR analysis (cDNA FLS), but attempts to join exon 24 to adjacent exons failed. RT-PCR analysis was useful in joining several regions of the gene together into cDNAs, but is limited in the ability to obtain certain regions of a cDNA.

Exon 11 has not been obtained in either the genomic sequence, an RT product, or a cDNA in humans. Several RT-PCRs were attempted using a variety of tissues, but no products were obtained that contained this exon. Primers, based on the rat sequence, were designed to sequence out from this exon using either cosmid DNA or genomic DNA as a template, but no readable sequence was obtained. PCR was tried using primers designed to the intron immediately flanking the suspected gap in sequence. One PCR product was obtained that was larger than the size predicted from the genomic sequence. This product was subcloned, and sequenced. However, no readable sequence was obtained for this piece of DNA. Exon 11 has not been confirmed as being present in the



human copy of *SHANK3* and several efforts to obtain this exon have failed.

Other regions of difficult PCR include exons 22, 23 and 26. No RT-PCR product was obtained that contains exon 22. This exon was identified by A. Wong (1998) as part of a human cDNA from fetal brain tissue but it is not part of the rat cDNA. Exon 23 is extremely large in size, approximately 2kb, and is very G/C rich. A cDNA that contains this exon in its entirety has not been obtained. However exon 23 is present in the rat cDNA sequence and RT-PCR products have joined this exon to other exons of *SHANK3* indicating that it is part of the transcript. Exon 23 contains most of the proline rich region of the SHANK3 protein, and includes the sequence for both the Homer binding domain and the SH3 binding domain. Exon 26 has only been predicted, no human cDNA or RT product have been obtained for this exon. Exon 26 is homologous to the terminal sequence from the rat cDNA, and contains the sequence for the SAM domain. Both exon 23 and 26 are present in the genomic sequence and the rat cDNA sequence, but exon 26 has not yet been detected in a human cDNA.

Random Amplification of cDNA Ends (RACE) was used to try to confirm the transcriptional start site and 3' ends of *SHANK3*. Primers were designed to RACE 5' from exon 2, and exon 12 and 3' from exons 20, 23 and 26. Several sets of nested primers were designed to RACE adult heart, brain, muscle and kidney libraries. Different combinations of primers and PCR conditions were used, but no products were obtained.

Several cDNA libraries were also screened in hopes of obtaining a more complete cDNA of *SHANK3*. Libraries made from pancreas, colon cancer (CaCo) and retinal RNA were screened using several different regions of *SHANK3* as a probe. No positive clones



were obtained that contained any part of *SHANK3*.

#### Expression of *SHANK3*:

Northern Blot analysis was used to determine the spatial expression pattern of *SHANK3*. Multiple Tissue Northern Blots purchased from Clontech Laboratories Inc. were probed with a region of *SHANK3* (exons 12 through 16). Multiple transcripts were observed with varying levels of expression in the different tissues tested (figure 3-5). The 8kb transcript showed strong expression in both brain and heart with weaker expression in other tissues tested. The 10kb transcript was present in most tissues tested, but at very low levels. A 2kb transcript was observed with strong expression in heart and weaker expression in kidney and liver. A brain predominant transcript of 7.5kb was observed with some very weak expression in kidney and liver. The same probe was also hybridized to a Northern Blot containing RNA isolated from different regions of the brain (Clontech Laboratories Inc.). On this Northern Blot, strong expression of the 10kb transcript was observed in the cerebellum with lower levels in the other brain tissues tested. The 7.5kb and 8kb transcripts showed moderate expression in all tissues tested with the exception of a lower level of expression in the spinal cord and medulla. However, the 8kb transcript was absent from the cerebellum. The Northern blot analysis of *SHANK3* in humans shows that expression of this gene is complex with many different transcripts at different levels in different tissues.

#### Expression of *Shank3* in other organisms:

The cDNA of the rat orthologue of *SHANK3*, when used to probe a rat multiple tissue Northern Blot, revealed a single transcript approximately 7.5kb present only in





brain tissue (Boeckers, et al, 1999b). Since this differs from the human expression pattern in both the specificity and number of transcripts, the expression of *SHANK3* was investigated in other model organisms. A human cDNA probe (identical to the probe used for human Northern blot analysis, exons 12 to 16) was hybridized to multiple tissue Northern Blots made of RNA isolated from both mouse and pig tissues. Since the sequence of the human and rat gene are so similar, the hybridization was carried out at the same stringency used in the hybridization of the human Northern blots. The mouse Northern blot (made by P. Brinkman-Mills) contained RNA from brain, heart, lung, liver and kidney from adult mice. Two transcripts were observed on this blot (figure 3-6). The larger transcript, approximately 8kb, was found predominantly in brain with very low levels observed in lung. The smaller 7.5kb transcript was specific to the brain. The pig multiple tissue Northern blots contained RNA from many different tissues isolated from a neonatal pig, including several different regions of the brain, as well as brain, heart, liver, kidney and skeletal muscle isolated from pig embryos. *Shank3* expression in the pig was observed as a single band, approximately 7.5kb in size, that was present in several regions of brain, as well as a very faint band in lung (figure 3-7). Thus the expression of *Shank3* in these two model organisms is not as widespread as the expression in human tissues, and not as many transcripts are observed.

#### Characterization of Another Member of the SHANK Family of Proteins:

In a library screen of a fetal brain library using part of the *SHANK3* gene, A. Wong (1998) identified a cDNA that was similar to *SHANK3*, but not identical. The sequence of this cDNA was analyzed using computer web based software. A BLAST



search of the NCBI non-redundant database, identified this gene as *SSTRIP* (which was later identified in rat as *Shank1*). Thus a cDNA containing part of the 5' region of the human *SHANK1* gene was identified. An alignment of the cDNA sequence of this gene to the sequence of the human genome localized this gene to a region on chromosome 19q13.3.

*SHANK3* is a good candidate for the neurological abnormalities observed in the 22q13 deletion syndrome. *SHANK3* is highly expressed in brain tissue, and has a transcript that is predominant in the brain. The Shank family of proteins form part of the post-synaptic density that join the synaptic membrane receptors to the cytoskeleton and signal transduction machinery. Based on the known function of the Shank family of proteins and the expression pattern of *SHANK3*, this gene was further investigated as a good candidate gene for the neurological abnormalities of the 22q13 deletion syndrome.

### **Characterization of Patient Deletions:**

#### **Verifying the deletion of *SHANK3* in patients:**

If *SHANK3* is responsible for the neurological abnormalities associated with the 22q13 deletion syndrome then it will be deleted in all patients with the 22q13 deletion syndrome. The polymorphic minisatellite D22S163 (Armour and Jeffreys, 1991) is located within intron 7 of *SHANK3* (see figure 3-2) and allows for easy verification of the deletion of *SHANK3*. Patients' DNA and parental DNA, if available, were digested with *Sau3A*, electrophoresed and transferred to Southern blots (assisted by H. Cheung and



W.Y. Tse). These blots were then probed with the polymorphic region of the probe D22S163. In all but one of the cases examined, the patient showed a single band at this locus indicating either hemizyosity or homozygosity of this locus (table 3-2). The one exception, patient MRS, was heterozygous for this probe (analyzed below in Characterization of Unusual Patients). If parental DNA was available (72% of patients had both parent's DNA available, 41/57 patients), then confirmation of a deletion could usually be made by comparing the child's banding pattern to that of the parents (figure 3-8). If the child only had one band which only matched a band observed in the mother but not the father then a paternal germline deletion of this probe is present, and vice versa for a maternal germline deletion. However, if the child's only band matched a band in both the mother and the father then the probe was uninformative and no further information could be obtained. Approximately 34% of patients with parental DNA samples were uninformative for this locus (12/35 patients). If this minisatellite probe was uninformative then the deletion of *SHANK3* was verified using FISH analysis.

FISH analysis, using one of the cosmids in the region of the NT deletion, was used on patients with an uninformative minisatellite D22S163 to verify the deletion of *SHANK3*. The cosmids c202, cN1G3, cN85A3 and cN66C4 were used in the analysis on various patients. Results were obtained for a total of 16 patients in FISH analysis with one of these cosmids. These cosmid probes were tried on 10 other patient metaphase spreads, but did not yield results either due to poor metaphase cell preparations or poor probe labeling/hybridization. The remaining patients that were uninformative for the minisatellite D22S163 (13) had no available metaphase cell preparations and thus the



deletion of *SHANK3* could not be confirmed. The deletion of *SHANK3* has been confirmed in 33 out of 56 patients (59%).

#### Determining Deletion Size Using Microsatellites:

In order to obtain an approximate idea of the size of the deletion microsatellite primers were labeled and PCR analysis was carried out on the patient DNA and the parental DNA if available. The PCR reactions were run on a polyacrylamide gel to obtain resolution of a few base pairs (figure 3-9). The microsatellite primer pairs selected for this analysis were spread throughout the region known to be deleted in a small number of previously characterized patients (Nesslinger, et al, 1994), as well as several primer pairs located more proximal than the largest previously determined deletion, patient 49/NS.

Analysis of these microsatellites was carried out in a similar manner to the analysis of the minisatellite D22S163. If the child had only one band, then parental samples were necessary to confirm the presence of a deletion, whereas two alleles proved no deletion was present (figure 3-10). If parental DNA samples were unavailable, then observation of a single band at any given microsatellite locus is inconclusive as the child could be either homozygous or hemizygous. The results of the microsatellite analysis are summarized in table 3-2.

Not every primer pair was used on every individual. Once the patient showed a heterozygous locus only a few primer pairs proximal to this locus were used to confirm the upper limit of the region containing the breakpoint. In the region that showed a deletion, not all primer pairs were used: if two microsatellites showed the same deletion, then the primer pairs in between were not used.





Some of the microsatellite loci were more polymorphic than the others. The number of uninformative patients, where parental DNA was available, ranged from 0% (D22S272) to 59% (D22S1159) for the microsatellites (table 3-2). Many of the microsatellites were uninformative in 30% or less of the patients. Some patients showed large regions of uninformative microsatellites, such as patient 19. Further analysis of these uninformative patients was carried out by fluorescent in situ hybridization (FISH).

#### Determining Deletion Size Using FISH:

Cosmids throughout the region were labeled with biotin and used for Fluorescent *In Situ* Hybridization on patients with large stretches of uninformative microsatellites, patients without parental DNA samples, and the patients with smaller deletions. The biotinylated cosmids were detected using fluorescently labeled antibodies. The locations of these cosmids were available on the Sanger Center's web site ([www.sanger.ac.uk](http://www.sanger.ac.uk)). The selection of cosmids to be used on the patient samples was based on the results from the microsatellite data.

For simple deletions and unbalanced translocations leading to the deletion of 22q13, metaphase cells were examined for a single set of fluorescent signal on a small acrocentric chromosome to determine that a deletion was present (figure 3-11). Since the metaphase spreads were not banded the presence of the deleted chromosome 22 was determined by counting the number of chromosomes observed in the preparation. If the number of chromosomes observed was 46 and only one fluorescent signal was observed, then the patient was deleted for this particular cosmid. If two sets of fluorescent signals were observed on small acrocentric chromosomes then the patient was not deleted for this



particular cosmid (summarized in table 3-3). FISH analysis was successfully carried out on 29/41 patients with metaphase cell preparations available. In four of the patients with metaphase cell preparations available, FISH analysis did not work (poor metaphase preparations or probe). FISH analysis was not carried out on eight patients since no cosmids in the region would have contributed any information not obtained by the microsatellites.

Unfortunately metaphase cell preparations were not available for all patients (16/57; table 3-3). The patients without metaphase cell preparations did not have the deletion of *SHANK3* confirmed if the minisatellite was uninformative, and did not have the breakpoint containing region narrowed down. Patient 19, had a large region of uninformative microsatellites and no metaphase cell preparation, therefore the breakpoint of this patient is localized to a region 4.5Mb in size that cannot be narrowed down any further. Patient 55's breakpoint is localized to a region almost 3Mb in size and cannot be narrowed down any further, due to the unavailability of a metaphase cell preparation. A deletion in patients 7, 23, and 48 have not been confirmed since neither parental DNA or metaphase cell preparations are available. All of these patients have several adjacent microsatellite loci that show a single band, but without metaphase cell preparations the presence and the size of the deletion cannot be unequivocally established. Metaphase cell preparations were available on most of the remaining patients to allow for confirmation of a deletion and to refine the size of the deletion.

By combining the minisatellite, microsatellite and FISH analysis, we have obtained the approximate sizes of the deletions found in the 22q13 deletion syndrome in



56 patients (summarized in table 3-4 and figure3-12). These deletions show a wide range of size from the 130kb of the NT deletion to the >10Mb found in 54 and 55's deletions.

#### Determining Parent of Origin of Deletion:

Using the information obtained from the minisatellite and microsatellite analysis, the parent of origin of the deletions could be determined if parental DNA samples were available. Both paternal and maternal germline deletions were observed in these patients. There were considerably more paternal germline deletions (27/40) than maternal germline deletions (13/40) observed. A Chi square analysis with the null hypothesis that both types of deletions should occur in equal frequency yielded a Chi square value of 4.9 and a p-value of 0.025. Since the p-value is less than 0.5, there are significantly more paternal germline deletions than maternal germline deletions in this patient sample.

#### Analysis of the Phenotype of the 22q13 Deletion Syndrome Patients:

Phenotypic information on these patients was obtained through a variety of sources. Patients 2, 6, 9, 12, 14, 16, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 31, 32, 34, 35, 38, 39, 40, 41, 42, 44, 46, 47, 48, 49, 50, 52, 53, 54, 55 and JaA were all examined by the same group of clinicians at the parent's support group meeting for this syndrome. The phenotypic information of these patients was summarized and analyzed by Phelan, et al, 2001. A summary of this phenotypic data was obtained from G. Stapleton (Appendix B). Clinical information on the remaining patients was obtained through various medical records as well as through communication with the clinicians, doctors and parents of these children. The key clinical features of all the patients are summarized in table 3-5. Information on all of the features was not available for all of the patients, leaving several



gaps in the analysis. All patients were described as having developmental delays and delay or absence of expressive speech. However while detailed data on the extent of the developmental delay was available for the patients analyzed by G. Stapleton, many of the other patients were just described as having developmental delay. Most of the patients also presented with hypotonia, and several other mild dysmorphic features.

To determine if the size of the deletion observed in the patients has any effect on the clinical features observed, linear regression was carried out to obtain a coefficient of correlation (a sample result is shown in Appendix C). As summarized in table 3-6, the coefficient of correlation ranges between 0 and 0.6. A coefficient of +1 means there is a strong positive correlation, whereas a coefficient of -1 means that there is a strong negative correlation. However, if the coefficient of correlation is 0, then no correlation is observed between the two variables. The F-value was examined to determine if the correlation was significant or not. A F-value of 0.05 or smaller showed that the correlation was statistically significant. Several different clinical features were found to have a correlation to the size of the deletion. Many different components of the IQ as well as several of the other clinical features, including abnormalities of the mouth and jaw.

The clinical features were also compared to the parent of origin of the deletion to determine if there is any imprinting effect. This analysis was carried out using contingency tables and Chi square analysis to see if the two variables are independent (a sample result is shown in appendix D). The results of the Chi square analysis are summarized in table 3-7. In some of the Chi Square analysis the null hypothesis that the





two variables are independent of each other was rejected on a p-value of less than 0.05. If the null hypothesis was rejected, then there is the possibility that there is some effect on the parent of origin on the clinical feature. A few of the components of the IQ tests rejected the null hypothesis as well as a few dysmorphic features.

### **Characterization of Unusual Patients:**

#### **Patients with a Bisatellited Chromosome 22:**

Patients 4, 20, and 32 were observed to have a bisatellited chromosome 22 by cytogenetic analysis from the referring cytogeneticists (figure 3-13). Patient 4's metaphase spreads were examined for the origin of the centromeric and pericentromeric region by T. Babineau from the Department of Human Genetics, McGill University (personal communication). Both of the centromeric regions on this patient's chromosome were shown to be from chromosome 22 material. This chromosome was NOR positive and DA/DAPI negative. NOR positive means that there is an active nucleolar organizing region present on this bisatellited chromosome. DA/DAPI is a stain specific for the satellite of chromosome 15. A chromosome 22 specific paint showed that the entire chromosome is composed of 22 material. A chromosome 22 centromeric FISH probe, D22Z4, identified signals on either end of the bisatellited chromosome 22 indicating a duplication of the centromeric region (figure 3-13). Patient 4 presents with the neurological features commonly seen in the 22q13 deletion syndrome patients although is lacking hypotonia. This patient also has some features, such as preauricular pits, that are commonly associated with cat eye syndrome (CES). The other two patients, 20 and 32, showed a similar cytogenetic structure of a bisatellited chromosome 22.



The amount of the material duplicated on these bisatellited chromosomes was of interest to determine if any part of the CES critical region was duplicated and therefore contributing to the phenotype. These patients were analyzed by FISH using BACs and PACs from the CES critical region, PAC109L3, BAC238M15 and BAC609C6 (Footz, et al, 2001). These three genomic clones cover most of the CES critical region and contain the putative candidate genes for the CES phenotype. Since these unusual chromosomes are bisatellited, if a duplication of the CES critical region was observed then fluorescent signal would be observed at both ends of the bisatellited chromosome. No duplications of any of these probes were observed for any of these three patients. However terminal deletions were confirmed in all of these patients.

#### Patient with TAPVR:

Another patient, 36, was found to have totally anomalous pulmonary venous return (TAPVR), a uncommon heart malformation that is often associated with CES. Cytogenetically this patient was found to have a visible deletion of 22q13, but no indication of a duplication of 22q11. In order to verify the cytogenetics, the CES FISH probes listed above were used to determine if any part of the CES critical region was duplicated in this patient. A terminal deletion was confirmed in this patient, however, no duplication of the CES critical region was detected.

#### Translocation Patient (t14;22):

Patient 2 in this thesis was cytogenetically identified as having a 22q13 deletion as the result of a translocation between chromosome 22 and 14. This translocation is a familial balanced translocation that is found in the mother, maternal grandmother, a



maternal aunt and a maternal cousin. Patient 2 was analyzed with microsatellites and showed no deletion, indicating a possible microdeletion of 22q13. FISH was carried out with probes of the two cosmids that contain the *SHANK3* gene. The more proximal probe, cN66C4, was present in two copies, although one of these signals was consistently fainter than the other (figure 3-14). Since one signal is fainter than the other, part of this cosmid may be deleted on one of the chromosome 22s. The second probe show only a single copy indicating that a deletion of 22q13 is present in this patient and that the breakpoint may be just proximal to *SHANK3*.

#### Mosaic Patient (JaA):

Patient JaA was cytogenetically identified as having a terminal deletion of 22q13. Further examination of this patient's chromosomes revealed a second cell line with an apparently normal chromosome complement (Phelan, et al, 2001). Cytogenetic analysis of JaA's chromosomes showed that 30% of the metaphase spreads showed a deletion, whereas, 70% of the metaphase spreads were normal. JaA has severe developmental delay, hypotonia, and the typical facial dysmorphisms observed with the 22q13 deletion syndrome (table 3-5). Since this patient is mosaic with a normal complement of chromosomes, all of the informative microsatellites tested showed two alleles, indicative of no deletion.

#### Interstitial Deletion (MRS):

Patient MRS was characterized cytogenetically as having an interstitial deletion of 22q13 (L. Artifoni, personal communication). Analysis of this patient and parental DNA with the minisatellite probe D22S163 showed that this probe was not deleted in this



individual. However through further microsatellite analysis a maternal germline interstitial deletion was detected. The distal breakpoint in MRS is between D22S1141 and D22S274 (figure 3-12). This breakpoint is located in the interval 5389kb to 5840kb from the subtelomeric repeats. Thus this deletion does not overlap the 19 smallest terminal deletions (figure 3-12). MRS's proximal breakpoint was characterized by FISH analysis and located between the probes cN17H6 and D22S276, representing an uninformative region of 516kb. The deletion in patient MRS is a maternal germline deletion that covers approximately 3Mb from 22q13. Some of the larger terminal deletions of 22q13 may contain the region deleted in the interstitial deletion MRS, as well as the terminal region of 22q. MRS shows some clinical features of the 22q13 terminal deletion syndrome such as severe developmental delay and delay of expressive speech. MRS however does differ in clinical features from the 22q13 deletion syndrome such as: no hypotonia, poor response to simple commands (although hearing is normal), poor balance, no notable facial dysmorphisms, and cognitive development corresponding to sensomotor phase. MRS is an unusual patient that was identified as having an interstitial deletion of 22q13 that does not overlap with the critical region for the 22q13 deletion syndrome, and may represent a clinically distinct, but overlapping genetic syndrome.

Several unusual patients were characterized in this study. Three cases with a bisatellite, dicentric chromosome were observed, as well as a case of mosaicism, and an interstitial deletion.





Figure 3-1: Alignment of Human *SHANK3* with Rat *ProSAP2* at the nucleic acid level. Mismatched bases are highlighted in grey. Gaps in the sequence are identified by -. Domains identified by a box.

Consensus	CCGGGGCCGATGCAGCTAAACCGTGCCGCCG-CGCCGCCGCCGC-GC-CCTGCGGAG	57
SHANK3	CCGGGGCCGATGCAGCTAAACCGTGCCGCCGCGCCGCCGCCGCCGCCCTGCGGAG	57
ProSAP2	CCGGGGCCGATGCAGCTAAACCGTGCCGCCGTGCCGCCGCCCGCTGCGCCTGCGGAG	57
Consensus	CCCCCGAGCCGCTGTCCCCCGCGC-GGCCCCGGCCCCGGCCCCCCCCCGCCCCCTC	114
SHANK3	CCCCCGAGCCGCTGTCCCCCGCGCCGCCGCCGCCGCCGCCGCCGCCGCCCTC	114
ProSAP2	CCCCCGAGCCGCTGTCCCCCGCGCTGGCCCCGGCCCCGGCCCCCCCCCGCCCCCTC	114
Consensus	CCGCG-AGCGCGG-CG-GCGGG-CTC-GGCGGGGG--AGGGGGGCC-GGCGCGCG	171
SHANK3	CCGCGCAGCGCGCCGAGCGGGCTCCGCGGGGGGGAGGGGGGGCCGGGGCGCGG	171
ProSAP2	CCGCGTAGCGCGGTGCGGGACTCTGGCGGGGGTCAGGGGGGGCCAGGCGCGCG	170
Consensus	CGCG-G-AGTCCCCG-GCGCTCC--TC-CCG-CG-GA-C-G-C--GGCCC-GGC-C-	228
SHANK3	CGCGCGGAGTCCCGGGCGCTCCGTTCCCGGCGCGAGCGCCCGGGCCCGGGCCCG	228
ProSAP2	CGCG-G-AGTCCCCGTGCGCTCTCTCTCCGCGGGAACAGTCAGGGCCCCGGCGCT	225
Consensus	-GC-C-GGGATGGACGGCCCCGGGGCGAGCGCCGTGGTCTGCGCGTCGGCATCCCC	285
SHANK3	GGCGCGGGATGGACGGCCCCGGGGCCAGCGCCGTGGTCTGCGCGTCGGCATCCCC	285
ProSAP2	AGCACCGGATGGACGGCCCCGGGGCCAGCGCCGTGGTCTGCGCGTCGGCATCCCC	282
Consensus	GACCTGCA-CA-ACGAAGTGCTGCG-CTGGA-CC--CCGCGCCCGTGTGGGCGGCC	342
SHANK3	GACCTGCAGCAGACGAAGTGCTGCGCCTGGACCCGCGCCCGCTGTGGGCGGCC	342
ProSAP2	GACCTGCAACAACGAAGTGCTGCGCTCTGGATCCAACCGCGCCCGTGTGGGCGGCC	339
Consensus	AAGCAGCG-GTGCTCTGCGCCCTCAACCCAGCGCT-CAGGACGCGCTCAACTA-GGG	399
SHANK3	AAGCAGCGGCTGCTCTGCGCCCTCAACCCAGCGCTCCAGGACGCGCTCAACTATGG	399
ProSAP2	AAGCAGCGTGTGCTCTGCGCCCTCAACCCAGCGCTTCAGGACGCGCTCAACTACGG	396
Consensus	CT-TTCCAGCC-CCCTCCCGGG-CGCGCCGGCAAGTTCTTGGATGA-GAGCGGCTC	456
SHANK3	CTTTTCCAGCCGCCCTCCCGGGCCCGCGCCGCAAGTTCTTGGATGAGGAGCGGCTC	456
ProSAP2	CTATTCCAGCCTCCCTCCCGGGTCTGCGCCGCGAAGTTCTTGGATGAAGAGCGGCTC	453
Consensus	-T-CAGGA-TACCCGCC-AACCTGGACACGCCCTGCCCCA-CTGGAGTTTCGATAC	513
SHANK3	CTGCAGGAGTACCCGCCCAACCTGGACACGCCCTGCCCTACCTGGAGTTTCGATAC	513
ProSAP2	TTACAGGACTACCCGCCAACCTGGACACGCCCTGCCCTATCTGGAGTTTCGATAC	510
Consensus	AAGCGG-GAGTTTATGCCCAGAACCTCAT-GATGA-AAGCAGTTTGCAAAGCT-CAC	570
SHANK3	AAGCGGGCAGTTTATGCCCAGAACCTCATGATGATAAGCAGTTTGCAAAGCTTAC	570
ProSAP2	AAGCGGAGAGTTTATGCCCAGAACCTCATAGATGACACAGCAGTTTGCAAAGCTGAC	567
Consensus	ACAAAGGC-AACCTGAAGAAGTTCATGGACTA-GTCCAGCT-CA-AGCAC-GACAAG	627
SHANK3	ACAAAGGCGAACCTGAAGAAGTTCATGGACTAGTCCAGCTGCATAGCAGCAGACAAG	627
ProSAP2	ACAAAGGCAACCTGAAGAAGTTCATGGACTATGTCCAGCTACACAGCAGACACAAG	624
Consensus	GTGGC-CGCCTG-TGGACAAGGGGCTGGACCCAA-TTCCATGACCCTGACTCAGGA	684
SHANK3	GTGGCAGCCTGTGTGACAAGGGGCTGGACCCCAACTTCCATGACCCTGACTCAGGA	684
ProSAP2	GTGGCCCGCTGTGACAAGGGGCTGGACCCCAATTTCATGACCCTGACTCAGGA	681
Consensus	GAGTGCCC-CTGAGCCT-GCAGC-CAG-TGGACAACGCCAC-GACCT-CT-AAGGT-	741
SHANK3	GAGTGCCCGCTGAGCCTGCGAGCCAGCTGGACAACGCCACGGACCTGTAAAGGTG	741
ProSAP2	GAGTGCCCTCTGAGCCTTGCAGCACAGTTGGACAACGCCACTGACCTCTGAAGGTT	738
Consensus	CT- <del>---</del> AATGG-GGTGC-CA-CTGGACTTCCG-AC-CG-GATGGGCT-AC-GC-GT-	798
SHANK3	CTGAAGAATGGTGGTGCCACCTGGACTTCCGCACTCGGATGGGTCACTGCGCGTG	798
ProSAP2	CTTCGCAATGGGGTGCTATCTGGACTTCCGAACCGAGATGGGCTAACCGCTGTG	795
Consensus	CACTG-GCCAC-CG-CAGCGGAATGCGG-AGCA-TGACGACCTGCTGGACCTGGGG	855
SHANK3	CACTGTGCCACACGACGCGGAATGCGGCACTGACGACCTGCTGGACCTGGGG	855
ProSAP2	CACTGCGCCACCCGACAGCGGAATGCGGGAGCATGACGACCTGCTGGACCTGGGG	852
Consensus	GCTTACCTGACTACAAGGACAGCCGCGG-TGAC-CCCCT-TACCA-AG-GCCCT-	912
SHANK3	GCTTACCTGACTACAAGGACAGCCGCGGCTTGACACCCCTCTACCAACGCGCCCTG	912
ProSAP2	GCTTACCTGACTACAAGGACAGCCGCGGCTGACGCCCTGTACCATAGTGCCTA	909
Consensus	GGGGG-GGGGATGCCCTCTGCTGTGAGCTGCTTCTCCA-GA-CACGC-CAG-TGGGG	969
SHANK3	GGGGGTGGGGATGCCCTCTGCTGTGAGCTGCTTCTCCACGACACGCTCAGCTGGGG	969
ProSAP2	GGGGGGGGGATGCCCTCTGCTGTGAGCTGCTTCTCCATGATCACGCACAGTTGGGG	966
Consensus	A-CAC-GACGAGAATGGCTGGCAGGAGATCCA-CAGGCTG-CGCTTTGGGCA-GT-	1026
SHANK3	ATACCGACGAGAATGGCTGGCAGGAGATCCACAGGCTGCGCTTTGGGCACGTG	1026
ProSAP2	ACCACTGACGAGAATGGCTGGCAGGAGATCCAACAGGCTGTCGCTTTGGGCATGTA	1023
Consensus	CAGCA-TGGAGCACCTGCTGTTCTATGGGGC--ACATGGG-GCCCGAAGCGCTCG	1083
SHANK3	CAGCATCTGGAGCACCTGCTGTTCTATGGGGCAGCATGGGGGCCAGAACGCGCTCG	1083
ProSAP2	CAGCACTGGAGCACCTGCTGTTCTATGGGGCAACATGGGTGCCGCAAGCGCTCG	1080

ANKYRIN REPEATS



Figure 3-1 continued: Alignment of Human *SHANK3* with Rat *ProSAP2* at the nucleic acid level. Mismatched bases are highlighted in grey. Gaps in the sequence are identified by -. Domains identified by a box.

Consensus	GG-AACACAGCC-TGCACATCTGTGCCCTCTA-AACCAGGAGAGCTGTGC-CG-GTC	1140
SHANK3	GGGAACACAGCCCTGCACATCTGTGCCCTCTACAACCAGGAGAGCTGTGCTGTGTC	1140
ProSAP2	GGAAACACAGCCCTGCACATCTGTGCCCTCTATAAACCAGGAGAGCTGTGCCCGCTC	1137
Consensus	CTGCT-TTCCGTGG-GC-AACA-GGA-GTCCGCAA-TACAACAGCCAGACAGCCTTC	1197
SHANK3	CTGCTCTTCCGTGGAGCTAACAGGGATGTCCGCAACTACAACAGCCAGACAGCCTTC	1197
ProSAP2	CTGCTTTTCCGTGTGCCAACAGGACGTCCGCAATTACAACAGCCAGACAGCCTTC	1194
Consensus	CAGGTGGCCAT-AT-GCAGGGAACTTTGAGCTTGC-GAGGT-ATCAAGACCCACAAA	1254
SHANK3	CAGGTGGCCATCATCGCAGGGAACTTTGAGCTTGCAGAGGTTATCAAGACCCACAAA	1254
ProSAP2	CAGGTGGCCATTATTGACAGGGAACTTTGAGCTTGCAGGTAATCAAGACCCACAAA	1251
Consensus	GACTCGGATGT-GTACCATTGAGGAAACCCAGCTATGC-AAGCG-CG-CG-CTG	1311
SHANK3	GACTCGGATGTTGTATACATTAGGGAACCCAGCTATGCGAAGCGGCGGCGACTG	1311
ProSAP2	GACTCGGATGTCTATACATTAGGGAACCCAGCTATGCAAGCGACGACGTCTG	1308
Consensus	GCTGGCCC-AGTGGCTTGGCATCCCTCGGCC--T-CAGCGCTCAGCCAG-GATATC	1368
SHANK3	GCTGGCCCAGTGGCTTGGCATCCCTCGGCCCTGACGCGCTCAGCCAGCGATATC	1368
ProSAP2	GCTGGCCCAAGTGGCTTGGCATCCCTCGGCCCTTACAGCGCTCAGCCAGTATATC	1365
Consensus	AACCTGAAGGG-GA---C---CAGC--CT-C---GG-CCC-C-CT--GAAGCCTC	1425
SHANK3	AACCTGAAGGGGAGGACAGCCAGCAGCTTCTCCTGGACCTCGCTGAGAAGCCTC	1425
ProSAP2	AACTGAAGGGTGACAGCCCGCAGCTTCTCC--CGGGCCCACTCTCCGAAGCCTC	1419
Consensus	CC-CACCA-CTGCTGCTCCAG-GGCT-CA-GAGGAGAAAGA-CG-GAC-GGGATG--	1482
SHANK3	CCCCACCACTGCTGCTCCAGCGGCTGCAAGAGGAGAAAGATCGTACCGGGATGCC	1482
ProSAP2	CCTCACCAACTGCTGCTCCAGAGGCTTCAGGAGGAGAAAGACCGGACAGGGATGGT	1476
Consensus	GA-CAGGAGA-C-ACATCAG-GG-CC-T-AGCAGGCAGGG-CGGCCA-AGCAAGATC	1539
SHANK3	GACCCAGGAGACCAACATCAGTGGCCCTTTAGCAGGCAGGGCCGGCCAAGCAAGATC	1539
ProSAP2	GAGCAGGAGAACACATCAGCGGTCCCTCAGCAGGCAGGGCGGCCACAGCAAGATC	1533
Consensus	AGCCCCAGCGGGCCCGCGGATCCGGCCCCGCGCCCGCCCGCCCGCGCTCTCCC	1596
SHANK3	AGCCCCAGCGGGCCCGCGGATCCGGCCCCGCGCCCGCCCGCCCGCGCTCTCCC	1596
ProSAP2	AGCCCCAGCGGGCCCGCGGATCCGGCCCCGCGCCCGCCCGCCCGCGCTCTCCC	1590
Consensus	GCGCCCCCGCG-CCGCCGCCCGGGGCCGAAGCGGAACTTTACAG-GCCCTCCCC	1653
SHANK3	GCGCCCCCGCGCCCGGCCCGGGGCCGAAGCGGAACTTTACAGCGCCCTCCCC	1653
ProSAP2	GCGCCCCCGCGCCCGGCCCGGGGCCGAAGCGGAACTTTACAGTGCCTCCCC	1647
Consensus	GGCCGCAAGTTTCATCGC-GTGAAGGCGCACAGCCCGCAGGG-GA-GGCGAGATCCCG	1710
SHANK3	GGCCGCAAGTTTCATCGCGTGAAGGCGCACAGCCCGCAGGGTGAAGGCGAGATCCCG	1710
ProSAP2	GGCCGCAAGTTTCATCGCTGTGAAGGCGCACAGCCCGCAGGGCGAGGGCGAGATCCCG	1704
Consensus	CTGCACCCGCGCGAGGCGGTGAAGGTGCTCAGCATTTGGGGAGGGCGGTTTCTGGGAG	1767
SHANK3	CTGCACCCGCGCGAGGCGGTGAAGGTGCTCAGCATTTGGGGAGGGCGGTTTCTGGGAG	1767
ProSAP2	CTGCACCCGCGCGAGGCGGTGAAGGTGCTCAGCATTTGGGGAGGGCGGTTTCTGGGAG	1761
Consensus	GGAACCGTGAA-GGCCG-AC-GGCTGGTTCCC-GC-GACTG-GTGAGGAAGTGCAG	1824
SHANK3	GGAACCGTGAAAGGCCGACGGCTGGTTCCCGCGACTGCGTGGAGGAAGTGCAG	1824
ProSAP2	GGAACCGTGAAAGGCCGTACAGGCTGGTTCCCGCTGACTGTGTGAGGAAGTGCAG	1818
Consensus	ATG-G-CAG-ATGACACACGGC-TGAAAC-CG-GAGGACCGACGAAGCG-CT-TT-	1881
SHANK3	ATGAGCAGCATGACACACGGCTGAAACCGGGAGGACCGGACGAAGCGCTCTTT	1881
ProSAP2	ATCGGACAGTATGACACACGGCATGAAACTCGAGAGGACCGGACGAAGCGCTCTTTT	1875
Consensus	CG-CACTACAC-GTGGG-TCCTA-GACAGCCTCAC-TCACACAG-GATTATGTCATT	1938
SHANK3	CGGCACTACACAGTGGGCTCCTACGACAGCTCACCTCACACAGGATTATGTCATT	1938
ProSAP2	CGGCACTACACTGTGGGTTCTTATGACAGCTCACTTCACACAGTATTATGTCATT	1932
Consensus	GATGA-AA-GTGGT-TCCGTGCA-AAACGGGACCA-GAGGG-TTTGG-TTTGT-CTC	1995
SHANK3	GATGACAAAGTGGCTTCTCTGCAGAAACGGGACCAAGAGGGCTTTGGTTTGTGCTC	1995
ProSAP2	GATGATAAGTGGCTATCTCTGCAAAACGGGACCATGAGGGTTTGGCTTTGTCTC	1989
Consensus	CGGGGAGCCAAAGCAGAGACCCCAT-GAGGAGTT-AC-CCCAC-CC-GCCTTCCC-	2052
SHANK3	CGGGGAGCCAAAGCAGAGACCCCATGAGGAGTTTACGCCACAGCCAGCCTTCCCG	2052
ProSAP2	CGGGGAGCCAAAGCAGAGACCCCATGAGGAGTTTACACCCACACTGCCTTCCCT	2046
Consensus	GCGCT-CAGTA-CT-GAGTC-GT-GA-GTGA-GGTGTGGCTTGA-GGC-GGGCT-	2109
SHANK3	GCGCTGCAGTATCTCGAGTCGTTGAGCTGGAGGGTGTGGCTTGAAGGCGGGCTC	2109
ProSAP2	GCGCTTCAGTACCTTGAGTCTGTAGATGTGAAGGTGTGGCTTGAAGGCTGGGCTT	2103
Consensus	CGCAC-GG-GACTTCCCTCAT-GAGGT-AACGG-GTGAACGT-GTGAAGGT-GGACAC	2166
SHANK3	CGCACGGGAGACTTCTCATCGAGGTGAACGGGTGAACGTGGTGAAGGTGCGACAC	2166
ProSAP2	CGCACTGGGACTTCTCATTTAGGTAAACGGAGTGAACGTGTAAGGTGGACAC	2160

SH3 DOMAIN

PDZ DOMAIN





Figure 3-1 continued: Alignment of Human *SHANK3* with Rat *ProSAP2* at the nucleic acid level. Mismatched bases are highlighted in grey. Gaps in the sequence are identified by -. Domains identified by a box.

Consensus	AAGCA-GTGGTG-TCT-AT-CG-CAGGGTGGCAACCG-CT-GTCATGAAGGTTGTG	2223
SHANK3	AAGCAGGTGGTGGCTCTGATTCGCCAGGGTGGCAACCGCTCGTCATGAAGGTTGTG	2223
ProSAP2	AAGCAAGTGGTGGGTCTCATCCGTCAGGGTGGCAACCGCTCTGTCATGAAGGTTGTG	2217
Consensus	TCTGT-AC-AGGAAGCCAGA-GAGGA--G-GCTCGGCGCAGAGCCCCACC-CC-CCC	2280
SHANK3	TCTGTGACAAGGAAGCCAGAAGAGGACGGGGCTCGGCGCAGAGCCCCACCGCCCCCC	2280
ProSAP2	TCTGTACCGGAACCCAGAGGAGGATAGTGCTCGGCGCAGAGCCCCACCACCTCCC	2274
Consensus	AAGAGGGCCCCAGCACCAC-CTGACCCTGCG-TCCAAGTCCATGAC-GCTGAGCTC	2337
SHANK3	AAGAGGGCCCCAGCACCACACTGACCCTGCGCTCCAAGTCCATGACAGCTGAGCTC	2337
ProSAP2	AAGAGGGCCCCAGCACCACGCTGACCCTGCGGTCCAAGTCCATGACGGCTGAGCTC	2331
Consensus	GAGGA-CT-GC-TCCATTCCGAGAAG-AAAGGGGAGAAG-TGGA-GAGAT-CTGGC-	2394
SHANK3	GAGGAACCTGGCTTCCATTTCGGAAGAAAGGGGAGAAGCTGGACGAGATGCTGGCA	2394
ProSAP2	GAGGAACCTGCTTCCATTTCGGAAGAAAGGGGAGAAGTTGGATGAGATCCTGGCG	2388
Consensus	G--GC-GC-GA-CCAACGCTG-GGCCAGACAT-GCAGACGC-GA-TCCAG-GC-GCC	2451
SHANK3	GCCCGCCGAGAGCCAAACGCTGCGGCCAGACATCGCAGACGCAGACTCCAGAGGCCGC	2451
ProSAP2	GTGCTGCGGAACCAACGCTGAGCGCAGACATTGCAGACGCTGATTCCAGGCGAGCC	2445
Consensus	AC-GTCAA-CAG-GGCCACCAG-CGGAGGAT-AC-CC-GCCGAGAT-AGCTCATTTG	2508
SHANK3	ACCGTCAAACAGAGGCCACCAGTCGGAGGATACACCCGCCGAGATTAGCTCATTTG	2508
ProSAP2	ACTGTCAAGCAGCGGCCACCAGCGCGAGGATTACCCCTGCCGAGATCAGCTCATTTG	2502
Consensus	TTTGA-CG-CAGGGCCTCCC-GGCCAGAGAAGCTGCCGGGCTC--TGCGGAAGGGG	2565
SHANK3	TTTGAACGCGCAGGGCTCCAGGCCAGAGAAGCTGCCGGGCTCCTTGCGGAAGGGG	2565
ProSAP2	TTTGAGCGCAGGGGCTCCCGGGCCAGAGAAGCTGCCGGGCTCTCTGCGGAAGGGG	2559
Consensus	ATTCCACGGACCAA-TCTGTAGGGGAGGA-GAGAAGCTGGC-TCCCT-CTGGAAGGG	2622
SHANK3	ATTCCACGGACCAAGTCTGTAGGGGAGGACGAGAAGCTGGCTCCCTGCTGAAGGG	2622
ProSAP2	ATTCCACGGACCAAAATCTGTAGGGGAGGATGAGAAGCTGGCATCCCTACTGGAAGGG	2616
Consensus	CG-TTCCC-CG-AGCAC-TC-ATGCAAGAC-C-GTGC-GA-GG-CG-GGCAT-CCG	2679
SHANK3	CGCTTCCCGCGGAGCACCTCGATGCAAGACCCGGTGCAGCGAGGGTCGCGGCATCCCG	2679
ProSAP2	CGTTTCCCACGACACATCAATGCAAGACACAGTGCCTGAAGGCGAGGCATTCGG	2673
Consensus	CCCCC-CCGACAGACGC-CCGCC-CCCC-CCCGCGCCTACTACTTGACTC-GGG	2736
SHANK3	CCCCCGCGCAGACCGCGCCCTCCCGCGCGCGCCTACTACTTGACTCGGGG	2736
ProSAP2	CCCCCACCGCAGACCGCCCGCCACCCCAACCCGCGCCTACTACTTGACTCCGGG	2730
Consensus	CC-CCCC--CCTTCTC-CC-CCGCC-CC-CC--GGGCGC-GCCTA-GACAC-GTG	2793
SHANK3	CCGCCCCCGCCTTCTCTGCGCGCCCGCCGCC--GGGCGCGCCTACGACACGGTG	2790
ProSAP2	CCACCCCCACCTTCTCAACACCGCCACCACCACCGGGCGGGCCTATGACACTGTG	2787
Consensus	CGCTCCAGCTTCAAGCC-GGCTTGGAGGC-CG-CTGGG-GC-GG-GC-GC-GGCCTG	2850
SHANK3	CGCTCCAGCTTCAAGCCCGGCTTGGAGGCGCGCTTGGGCGCGGGGCGCTGCGGCCCTG	2847
ProSAP2	CGCTCCAGCTTCAAGCCAGGCTTGGAGGCTGCTGCGGTGACGGGCGAGTGGCCTG	2844
Consensus	TA-GA--C-GGC-C--C-CT-GGCCCGCTGCC-TA-CC-GAGCG-CAGAAGCG-GC-	2907
SHANK3	TACGAGCCGGGGCGGCGCTTGGGCCGCTGCCGTATCCGAGCGGAGAGCGCGCG	2904
ProSAP2	TATGATTCTGGCACACCTCTGGGCCGCTGCCCTACCTAGAGCGCCAGAAGCGTGCA	2901
Consensus	CGCTCCATGATCAT--TGCAGGACTC-GCGCC-GA--GGGCGA-G--CC-CG-CC-	2964
SHANK3	CGCTCCATGATCATCTCTGAGGACTCGGCGCCGAGTCCGGCGACGCCCTCGACCC	2961
ProSAP2	CGCTCCATGATCATATTGAGGACTCTGCGCCAGAAGTGGCGATGTACCCGGCGCT	2958
Consensus	CGCC-GC-GCCAC-CCGCC-GAGCG-CCCAAGCGCGGCC-CGGCCG-C-GGCC-	3021
SHANK3	CCGCCCCGCGCCACCCGCCGAGCGACCAAGCGCGCGCGCGCGCCCGCGGCC	3018
ProSAP2	GCGCTGACGCACACCGCCTGAGCGCCCAAGCGCGCGCCTCGGCGCTGAGGCCCT	3015
Consensus	GA-AG-CCCTA-GCCAACCTGGGCGCCTTCAG-GCCAGCCTCTT-GCTCCGTC-AA-	3078
SHANK3	GACAGCCCTTACGCCCAACCTGGGCGCCTTCAGCGCCAGCCTCTTGGCTCCGTCAG	3075
ProSAP2	GATAGTCCCTATGCCAACCTGGGCGCCTTCAGTGCCAGCCTCTTGTCTCCGTGAAA	3072
Consensus	CCGACGCGCCGAAGAG-CC-CTGGTGAAGCAGCT-CAGGTGAGGACGC-CAGGAG	3135
SHANK3	CCGACGCGCCGAAGAGCCCTGGTGAAGCAGCTGAGGTGAGGACGCGCAGGAG	3132
ProSAP2	CCGACGCGCCGAAGAGTCCGCTGGTGAAGCAGCTTACGTTGAGGACGCTCAGGAG	3129
Consensus	CGCGCGCG--TGGCCGTGGG-AGCCC-GG-CC-G--GG-GG-AGCTT-GC-CG-GA-	3192
SHANK3	CGCGCGCGCCTGGCGTGGGACGCCCGCTCCGCGCGCGCGAGCTTGGCCCGGAG	3189
ProSAP2	CGCGCGCGCTTGGCCGTGGGTAGCCCGGACCACTGGGTGGAAGCTTGCACAGAA	3186
Consensus	CCCTCCCC-AC-CACCGCG-CC-CG-CCGGG-GGCCT-GACTAC-GC-C-GG-GA-	3249
SHANK3	CCCTCCCCGACCCACCGCGGTCGCGCCCGGGTGGCTCGACTACGGCGGGGCGAT	3246
ProSAP2	CCCTCCCCAACGACCGCGGGCCCGACCGGGCGGCTTACTACAGCTCTGAGAA	3243

HOMER BINDING Motif #1



Figure 3-1 continued: Alignment of Human *SHANK3* with Rat *ProSAP2* at the nucleic acid level. Mismatched bases are highlighted in grey. Gaps in the sequence are identified by -. Domains identified by a box.

Consensus	GGCC-GGGGCTC-C-TT-GGCGGCC--GCC--GCC-AG---A---GCGGC---G	3306
SHANK3	GGCCCGGGGCTCGCGTTCGGCGGCCCGGGCCCG-GCCAAG--GACCGCGCGCTGGAG	3300
ProSAP2	GGCCTGGGGCTCACCCTTGCGGCGCCCTAGCCCTGGCCAGTCAAGGAGCGCGCCTG	3300
Consensus	GAG--GCG-CGC---C---TGT-T-C-TGTC-GTGGG-GCCATCGAGGGCAGC-C-	3363
SHANK3	GAGCGCGCGCTCCACTG-TGT-TCC-TGTCGTGGGGCCATCGAGGGCAGCGCC	3354
ProSAP2	GAGGAGCGACCCGTTCCACTGTGTCTCTGTGTGGTGCCATCGAGGGCAGCCCT	3357
Consensus	CCC-GCGCGGATCTGCCATCCCTACAGCCCTCCCGCTCCAT-GA-GAGCGCTCCTG	3420
SHANK3	CCCGGCGGGATCTGCCATCCCTACAGCCCTCCCGCTCCATCGACGAGCGCTCCTG	3411
ProSAP2	CCCAGCGCGGATCTGCCATCCCTACAGCCCTCCCGCTCCATTGATGAGCGCTCCTG	3414
Consensus	GGGAC-GGC-CCACC-C-GGCCGCGA--TGCTGCT-CCCTCCCC-GT-TCTGC-CTG	3477
SHANK3	GGGACCGGGCCCCACCGCGGCGCGACCTGTGCTGCCCTCCCGGTGTCTGCGCTG	3468
ProSAP2	GAGCAGGCGCCACCACTGGCCGCGATTGTGCTGCTCCCTCCCTGTCTGTGCTGTG	3471
Consensus	AAGCC-TTGGTC-G-GG-CC-AGCCT-GGGCCCTC-GG-TCCACCTTCATCCA-CC-	3534
SHANK3	AAGCGTGTGGTCAGCGGGCCGAGCCTGGGGCCCTCGGGTTCACCTTCATCCACCA	3525
ProSAP2	AAGCCATTGGTCTGGTGGTCCGAGCCTTGGGCCCTCAGGCTCCACCTTCATCCATCT	3528
Consensus	CT-AC-GGCAAAACC-TGGA-CC-AGCTACCCCT-GC-CTTGC-CTGGCTGCCCGA	3591
SHANK3	CTACCCGGCAAAACCCTGGACCCAGCTCACCCCTGGCCCTTGCCCTGGGTGCCCGA	3582
ProSAP2	CTTACTGGCAAAACCCTTGGATCCTAGCTCACCCCTAGCTCTTGCTCTGGCTGCCCGA	3585
Consensus	GAGCG-GCTCTGGCCTC-CA--C-CC-TCCCGTCCCCCACACCCGTGCACAGTCC-	3648
SHANK3	GAGCGAGCTCTGGCCTCCAGGCGCCCTCCCGGTCCCCACACCCGTGCACAGTCC	3639
ProSAP2	GAGGGGCTCTGGCCTCGCAACACCTTCCCGGTCCCCCACACCCGTGCACAGTCT	3642
Consensus	GA-GC-GACCGCCC-GGACCCT-TTGTGGATGT-CA--CCCG-GAC-C-GAG-GA	3705
SHANK3	GACGCCGACCGCCCCGACCCTGTTTGTGGATGTACAGGCCGGGACCCAGAGCGA	3696
ProSAP2	GATGCTGACCGCCTGGACCCCTCTTGTGGATGTGCAACCCGAGAGATCCGAGAGA	3699
Consensus	GG--CC-TGGC-TCCCC-GC-TTCTCCCC-CGGAG-CCAGCCTGGATTCC-GTGCCT	3762
SHANK3	GGGTCCCTGGCTTCCCCGGCTTCTCCCCAGGAGCCAGCCTGGATTCCGTGCTCT	3753
ProSAP2	GGACCCCTGGCCTCCCCAGCCTTCTCCCCCGAGTCCAGCCTGGATTCCAGTGCCT	3756
Consensus	GCTCGCAG-GAGGCAGAGAAG--C-C-CGGGA-GAGCGGAAGTCACC-GAGGACAAG	3819
SHANK3	GCTCGAGGGAGGCAGAGAAGTCCCCGGGAGGAGCGGAAGTACCCGAGGACAAG	3810
ProSAP2	GCTCGCAGAGGCAGAGAAGCCCACTCGGGAAGAGCGGAAGTACACAGGAGACAAG	3813
Consensus	AA-TCCATGATCCTCAGCGTC-TGGACAC-TCC-TGCA-CGGCCAGCTGG-CTCAT-	3876
SHANK3	AAGTCCATGATCCTCAGCGTCTTGACACATCCCTGCAGCGGCAGCTGGCTCTATC	3867
ProSAP2	AAATCCATGATCCTCAGCGTCTTGACACGTCTTGCAACGGCCAGCTGGTCTCAT	3870
Consensus	GTGTGCA-GCCACCAGCAA-GG-CAGGAGCCCA-CAGGCTGGGGG--G--GA-GAG	3933
SHANK3	GTGTGCAAGCCACAGCAACGGGAGGAGCCAGCAGGCTGGGGGGGGCGAAGAG	3924
ProSAP2	GTGTGCAATGCCACCAGCAATGGACAGGAGCCCAACAGGCTGGGGGCTGAAGAGGAG	3927
Consensus	---C---CGGG-AC-CCGGAG-TGGCCCC--CCCCATGCAG-CAGC-GCTGTGGCA	3990
SHANK3	GAGCGCCCGGGCACCCTGGAGTTGGCCCCGGGCCCATGCAGTCAAGCGGCTGTGGCA	3981
ProSAP2	CGCC---CGGGTACTCCGAGCTGGCCCCAACCCCATGCAGGCAGCAGCTGTGGCA	3981
Consensus	GAGCCC-TGCC-AGCCC-CG-GCCAGCCCCCTGG--GCA-CCC-GCAGA--CCGGG	4047
SHANK3	GAGCCCCTGCCAGCCCCCGGCCAGCCCCCTGGTGGCACC CGGCAGAGCGCGG	4038
ProSAP2	GAGCCCATGCCAAGCCACGAGCCAGCCCCCTGGCAGCATCCAGCAGATCCCGGG	4038
Consensus	CCAGGCCA-GGCAGCTCAGAGGA-GAGCCAGAGCTGGT-TT-GCTGTGAACCTGCCA	4104
SHANK3	CCAGGCCAGGCGAGCTCAGAGGAAGAGCCAGAGCTGGTGTTCGCTGTGAACCTGCCA	4095
ProSAP2	CCAGGCCAAGGCAGCTCAGAGGAGGAGCCAGAGCTGGTATTGCTGTGTGAACCTGCCA	4095
Consensus	CCTGC-CAGCTGTC-TCCAGCGATGAGGAGACCAG-GAGGAGCTGGCCCG-ATTGGG	4161
SHANK3	CCTGCCAGCTGTCTCCAGCGATGAGGAGACCAGGAGGAGCTGGCCCGAATTGGG	4152
ProSAP2	CTCTGCTCAGCTGTCTCCAGCGATGAGGAGACCAGAGGAGCTGGCCCGCATGGG	4152
Consensus	T-TGTGCCACCCCTGAAGAGTTTGCCAA-GGG-TCCTGCTGGCCACCCCA-C-C-	4218
SHANK3	TTGTGTGCCACCCCTGAAGAGTTTGCCAACGGGGTCTGTGTGGCCACCCCACTCGCT	4209
ProSAP2	CTAGTGCCACCCCTGAAGAGTTTGCCATGGGATCTGTGTGGCCACCCCACTCGCA	4209
Consensus	GG-CCGGGCCCC--GCCACCACGCT-CCCAGCCCGGCCCTCAGGGAAGCCAGCAG-	4275
SHANK3	GGCCCGGGCCCTTCGCCACCACGGTGCCAGCCCGGCCCTCAGGGAAGCCAGCAGT	4266
ProSAP2	GGACGGGCCCCCTGCCACCACGCTACCCAGCCCGGCCCTCAGGGAAGCCAGCAGC	4266
Consensus	GAGC--CCCCCTGCCCTGAGTCTGCAGC-GACTCTGG-GT-GAGGAGGC-GACAC-	4332
SHANK3	GAGCCACCCCTGCCCTGAGTCTGCAGCCGACTCTGGGTGGAGGAGGCTGACACA	4323
ProSAP2	GAGTGCSCCCTGCCCTGAGTCTGCAGCTGACTCTGGAGTAGGAGGCGGACACT	4323

HOMER BINDING Motif #2





Figure 3-1 continued: Alignment of Human *SHANK3* with Rat *ProSAP2* at the nucleic acid level. Mismatched bases are highlighted in grey. Gaps in the sequence are identified by -. Domains identified by a box.

Consensus	CG-AGCTCCAG-GACCCCCACCTGGAGACCACAAGCACCAT-TCCAC-GTGTCACG	4389	
SHANK3	CGCAGCTCCACGACCCCCACCTGGAGACCACAAGCACCATCTCCACGGTGTCCAGC	4380	
ProSAP2	CGAAGCTCCAGTGACCCCCACCTGGAGACCACAAGCACCATTTCACAGTGTCCAGC	4380	
Consensus	ATGTCCACC-TGAGCTCGGAGAG-GG-GAACTCACTGACACCCACACCTCCTT-GC	4446	
SHANK3	ATGTCCACCTTGAGCTCGGAGAGCGGGGAACTCACTGACACCACACCTCCTTCGCT	4437	
ProSAP2	ATGTCCACCTTGAGCTCGGAGAGTGGAACTCACTGACACCCACACCTCCTTTGCC	4437	
Consensus	GA-GGACACACTTTTCTACTCGAGAAGCCACCACTGCCTCCGAGCCCAA-CTCAAG	4503	
SHANK3	GACGGACACACTTTTCTACTCGAGAAGCCACCACTGCCTCCGAGCCCAAAGCTCAAG	4494	
ProSAP2	GATGGACACACTTTTCTACTCGAGAAGCCACCACTGCCTCCGAGCCCAAAGCTCAAG	4494	SH3 BINDING Motif
Consensus	TCCCCGCTGGGGAAGGGGCGGTGACCTTCAGGAGCCCGCTGCTGAAGCA-TCCTCG	4560	
SHANK3	TCCCCGCTGGGGAAGGGGCGGTGACCTTCAGGAGCCCGCTGCTGAAGCAGTCTCTCG	4551	
ProSAP2	TCCCCGCTGGGGAAGGGGCGGTGACCTTCAGGAGCCCGCTGCTGAAGCAATCTCTCG	4551	
Consensus	GACAG-GAGCTCATGGCCAGCAGCACC-A-GCC-CCTCT-C-GGG-TG-C-TCTGC-	4617	
SHANK3	GACAGCAGCTCATGGCCAGCAGCACCAGCCGCTCTGCCGGGCTGGCCTCTGCTC	4608	
ProSAP2	GACAGTGAGCTCATGGCCAGCAGCACCATGCCACCTCTACTGGGTGACTTCTGCT	4608	
Consensus	GC-GGGCCTGCCGCCTCGCTACCTCTTCAGAGAAGGTCCAAGCT-TGGGGGGAC	4674	
SHANK3	GCCGGGCTGCCGCCTCGCTACCTCTTCAGAGAAGGTCCAAGCTATGGGGGGAC	4665	
ProSAP2	GCGGGGCTGCCGCCTCGCTACCTCTTCAGAGAAGGTCCAAGCTGTGGGGGGAC	4665	
Consensus	CCCCGTGGAGAG-CGGGGGCTCCCTGGGCTGA-GA-GACAAACCAACTGTGATCAGT	4731	
SHANK3	CCCCGTGGAGAGCCGGGGGCTCCCTGGGCTGAAGACGACAAACCAACTGTGATCAGT	4722	
ProSAP2	CCCCGTGGAGAGTCCGGGGCTCCCTGGGCTGAGGATGACAAACCAACTGTGATCAGT	4722	
Consensus	GAGCTCAGCTCCCG-CTGCAGCAGCTGAA-AA-GACAC-CG-TCC-TGGGGGAGGAA	4788	
SHANK3	GAGCTCAGCTCCCGCTGCAGCAGCTGAACAAGGACACGCTTCCCTGGGGGAGGAA	4779	
ProSAP2	GAGCTCAGCTCCCGCTGCAGCAGCTGAATAAGACACTCGCTTCTGGGGGAGGAA	4779	
Consensus	CCAGTTGTTGGCTGGGCAGCCTGCTGGACCTGC-AAGAAGTCGCCAT-GCAGCA	4845	
SHANK3	CCAGTTGTTGGCTGGGCAGCCTGCTGGACCTGCCAAGAAGTCGCCATGCGAGCA	4836	
ProSAP2	CCAGTTGTTGGCTGGGCAGCCTGCTGGACCTGCTAAGAAGTCGCCATGCGAGCA	4836	
Consensus	GCTCGGCTCTTCAGCAGCCTCGGTGAGCTGAGC-CCAT-TCAGCGCAGCGCAGCCCC	4902	
SHANK3	GCTCGGCTCTTCAGCAGCCTCGGTGAGCTGAGCTCCATTTACGCGCAGCGCAGCCCC	4893	
ProSAP2	GCTCGGCTCTTCAGCAGCCTCGGTGAGCTGAGCACCATCTCAGCGCAGCGCAGCCCC	4893	
Consensus	GGGGGCCCGGGCGG-GGGGCTC-TACTCGGTG-GGCCAG-GGCCG-TACCCCGTG	4959	
SHANK3	GGGGGCCCGGGCGCGGGGCTCTGCTACTCGGTGAGGCCAGTGGCCGTACCCCGTG	4950	
ProSAP2	GGGGGCCCGGGCGGAGGGGCTCTACTCGGTGCGGCCAGCGGCGGTACCCCGTG	4950	
Consensus	GCGAGACG-GCCCGAGCCC-GTGAA-CCCGC-TCGCTGGAGCGGTGGAGGGGCTG	5016	
SHANK3	GCGAGACGCGCCCGAGCCCGGTGAAGCCCGCTGCTGGAGCGGTGGAGGGGCTG	5007	
ProSAP2	GCGAGACGAGCCCGAGCCAGTGAAACCCGCATCGTGGAGCGGTGGAGGGGCTG	5007	
Consensus	GGGGCGGGCG-GGG-GGCG-GGGCGGCCCTTCGGCTCAGCC-CCCACCATCCTC	5073	
SHANK3	GGGGCGGGCGCGGGGGCGCAGGGCGGCCCTTCGGCTCAGCCGCCACCATCCTC	5064	
ProSAP2	GGGGCGGGCTGGGAGGCGCGGGCGGCCCTTCGGCTCAGCCCTCCACCATCCTC	5064	
Consensus	AAGTCGTCCAGCCTCTCCATCCCGCAGCA-CCCAAGGA-GTGCCTTCGTGGTGCG	5130	
SHANK3	AAGTCGTCCAGCCTCTCCATCCCGCAGCAGCCCAAGGAGTGCCTTCGTGGTGCG	5121	
ProSAP2	AAGTCGTCCAGCCTCTCCATCCCGCAGCAACCAAGGAAGTGCCTTCGTGGTGCG	5121	
Consensus	AG-GTGAG-GCGGCGAG-CGCTCCCCCTC-CC-TC-CCGCTGCCCTCGCC--C-C-	5187	
SHANK3	AGCGTGAGCGCGCGCAGTCGCTCCCCCTCGCGCTCGCGCTGCCCTCGCCCGCTGC	5178	
ProSAP2	AGTGTGAGTGCGGCGAGCGCTCCCCCTCACCATCTCCGCTGCCCTCGCCCTTCCT	5178	
Consensus	GGC-C-GGCCCC-G-GCC--CCCG-G-C--C----C--TTC-A-CAGAAGCC-CTG	5244	
SHANK3	GGCCCGGCCCCGGCGC--CCCG-GCCACGCCGACCTTCAGCAGAGAAGCCGCTG	5232	
ProSAP2	GGCTCTGGCCCCAGTGCCGGCCCGCTCGGC---CAATTCAA-CAGAAGCCCGCTG	5229	
Consensus	CAGCT-TGGAGCAAGTTCGA-GTGGGCGACTGGCTGGAGAGCATCCAC-TAGGCGAG	5301	
SHANK3	CAGCTTGGAGCAAGTTCGAGTGGGCGACTGGCTGGAGAGCATCCACCTAGGCGAG	5289	
ProSAP2	CAGCTTGGAGCAAGTTCGAGTGGGCGACTGGCTGGAGAGCATCCACTTAGGCGAG	5286	
Consensus	CACCG-GACCGCTTCGAGGACCATGAGAT-GAAGGCGC-CACCT-CC-GCGCT-ACC	5358	
SHANK3	CACCGGACCGCTTCGAGGACCATGAGATGAAGGCGCGCACTACCGCGCTTACC	5346	
ProSAP2	CACCGAGACCGCTTCGAGGACCATGAGATGAAGGCGCACACTGCTCGCGCTACCC	5343	SAM DOMAIN
Consensus	AAGGA-GACTTCGTGGAGCTGGG-GTCAC-CGCGT-GGCCACCGCATGAACATCGAG	5415	
SHANK3	AAGGACGACTTCGTGGAGCTGGGCGTCACGCGCTGGGCCACCGCATGAACATCGAG	5403	
ProSAP2	AAGGAAGACTTCGTGGAGCTGGGAGTCACGCGTGGGCCACCGCATGAACATCGAG	5400	



Figure 3-1 continued: Alignment of Human *SHANK3* with Rat *ProSAP2* at the nucleic acid level. Mismatched bases are highlighted in grey. Gaps in the sequence are identified by -. Domains identified by a box.

Consensus	CG-GCGCTCAGGCAGCTGGA-GGCAGCTGACGCCCC-C-CCC-CTCC-G--CC-GC-	5472
SHANK3	CGCGCGCTCAGGCAGCTGGACGGCAGCTGACGCCCCACCCCACTCCCGCCCGGCC	5460
ProSAP2	CGTGCGCTCAGGCAGCTGGATGGCAGCTGACGCCCTCTCCCTCTCCTGTTCTGCT	5457
Consensus	G-GCCCTGCCGGCAGGGCCCC--CCC-TACTCCAGGCCGACGCTCGGCTCGCCCC	5529
SHANK3	GTGCCCTGCCGGCAGGGCCCCCAACCCCTACTCCAGGCCGACGCTCGGCTCGCCCC	5517
ProSAP2	GCGCCCTGCCGGCAGGGCCCCCAACCC-TACTCCAGGCCGACGCTCGGCTCGCCCC	5513
Consensus	CTACCACGGCGCCCGGGCCAGGAATGTTGCATGAATCGTCCTGTTTGCTGTTGCTTG	5586
SHANK3	CTACCACGGCGCCCGGGCCAGGAATGTTGCATGAATCGTCCTGTTTGCTGTTGCTTG	5574
ProSAP2	CTACCACGGCGCCCGGGCCAGGAATGTTGCATGAATCGTCCTGTTTGCTGTTGCTTG	5570
Consensus	GAGACTTGCCCTGTACATTGCTTAGTGCCCTCCCTGCCGCGAACCACCCACCCAGCA	5643
SHANK3	GAGACTTGCCCTGTACATTGCTTAGTGCCCTCCCTGCCGCGAACCACCCACCCAGCA	5631
ProSAP2	GAGACTTGCCCTGTACATTGCTTAGTGCCCTCCCTGCCGCGAACCACCCACCCAGCA	5627
Consensus	CACAGTAAGGGCGCGGACCAGGGGGCTGGGTGGAAGGGGTTGGGGCAGGGTGCTC	5700
SHANK3	CACAGTAAGGGCGCGGACCAGGGGGCTGGGTGGAAGGGGTTGGGGCAGGGTGCTC	5688
ProSAP2	CACAGTAAGGGCGCGGACCAGGGGGCTGGGTGGAAGGGGTTGGGGCAGGGTGCTC	5684
Consensus	TGGCCTGACCACCTCCTCCACAGCTCCTGGTGGCCATTCTTCCAGAGGGGAACCTA	5757
SHANK3	TGGCCTGACCACCTCCTCCACAGCTCCTGGTGGCCATTCTTCCAGAGGGGAACCTA	5745
ProSAP2	TGGCCTGACCACCTCCTCCACAGCTCCTGGTGGCCATTCTTCCAGAGGGGAACCTA	5741
Consensus	GTCCAGCATGCGAGGTCAGGACACGCCTTGGTGACTCGGGGGAGGGGGAGACATT	5814
SHANK3	GTCCAGCATGCGAGGTCAGGACACGCCTTGGTGACTCGGGGGAGGGGGAGACATT	5802
ProSAP2	GTCCAGCATGCGAGGTCAGGACACGCCTTGGTGACTCGGGGGAGGGGGAGACATT	5798
Consensus	GGGG	5818
SHANK3	GGGG	5806
ProSAP2	GGGG	5802





Figure 3-2: A: Genomic organization of SHANK3. Each exon is represented by a vertical bar. A dotted vertical bar indicates a predicted exon. The genomic sequence is represented by the horizontal bar. Each of the protein interacting domains is identified. The breakpoint in the NT microdeletion is identified by an arrow. B: RT-PCR products obtained in this thesis. Vertical bars represent the exons obtained. Exons are joined by a solid line if confirmed in product, or a dotted line if not confirmed in the product.



**B:**

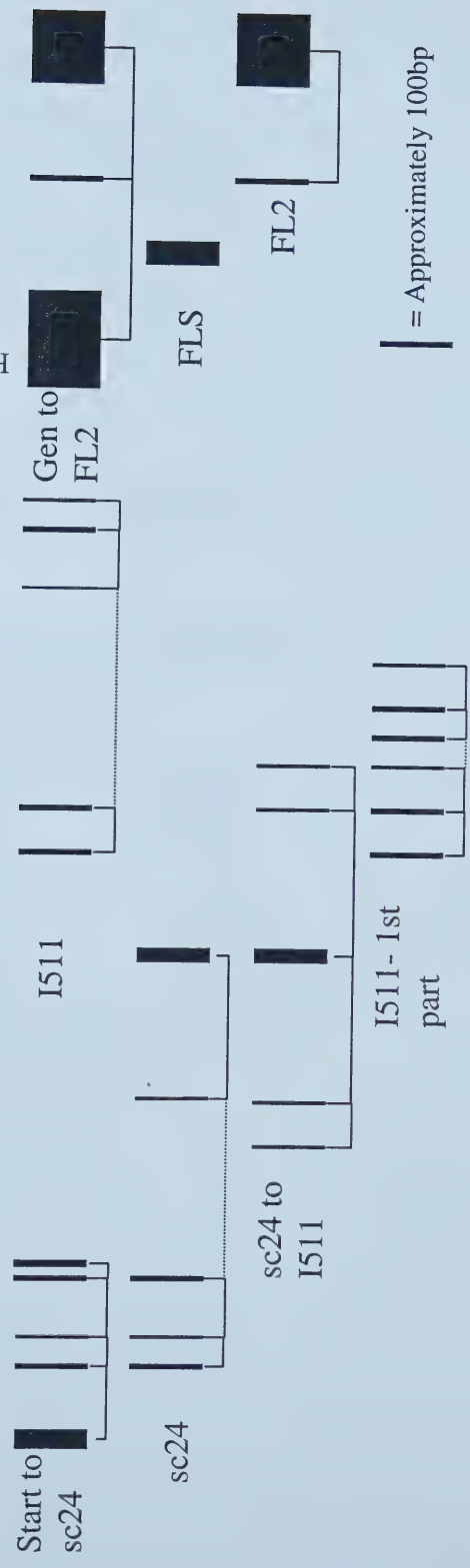




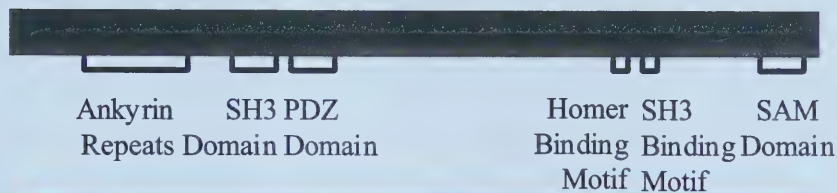




Figure 3-3: Similarity between SHANK3 and other Shank family members at the amino acid level. Proteins are represented by a horizontal bar and aligned based on the amino acid similarity to the human SHANK3. Key protein interacting domains are identified on SHANK3, and are found in the same order in the other proteins of this family. Amino acids are numbered across the top. Extremely high similarity at the amino acid level is observed between the rat and the human Shank3. High similarity is observed in the protein interacting domains between all members of the Shank family of proteins.



Human  
SHANK3



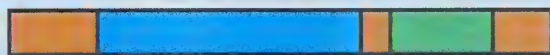
Rat Shank3



Rat Shank1a



Rat Shank2









-  = >90% similarity at the amino acid level
-  = 80 to 89% similarity at the amino acid level
-  = 60 to 79% similarity at the amino acid level
-  = 50 to 59% similarity at the amino acid level
-  = 30 to 49% similarity at the amino acid level
-  = 20 to 29% similarity at the amino acid level





Figure 3-4: Alignment of human SHANK3 domains with other Shank gene family members. Letters highlighted in black indicate identical amino acids, and letters highlighted in grey indicate conserved amino acids. (A) Ankyrin Repeat Domain; (B) PDZ Domain; (C) SAM Domain; (D) HOMER-Binding Domain; (E) SH3 Domain; (F) SH3-Binding Domain. The Ankyrin Repeat Domain is not found in Shank2, whereas all of the other domains are found in all three members of this family of proteins. An extremely high degree of similarity is found in these domains between all of the members of the Shank family.

(A)

Human SHANK3  
Rat Shank3  
Rat Shank1A

VOLHSTDKVARLLDKGLDPNFDPSGCEPLSLAAQLDNATDLKVLKNGGAHLDFRTRDGLTAVHCAFRORNAALTTVLIDIGASPDYKDSRGHTPLVYHSALGGGDALCCCELLLHDHAQLGTTD  
VOLHSTDKVARLLDKGLDPNFDPSGCEPLSLAAQLDNATDLKVLKNGGAHLDFRTRDGLTAVHCAFRORNAALTTVLIDIGASPDYKDSRGHTPLVYHSALGGGDALCCCELLLHDHAQLGTTD  
VOLSTSKVARLLDKGLDPNFDPSGCEPLSLAAQTEGSGVEVIRLQTEGGAHLDFRTRDGLTAVHCAFRORNAALTTVLIDIGASPDYKDSRGHTPLVYHSALGGGDALCCCELLLHDHAQLGTTD

Human SHANK3  
Rat Shank3  
Rat Shank1A

ENGWQETHOACRFHGVOHLEHLFYGADMGAGNAGSNVAHHICALYNQESCARVLFLFRGANRDVRNYSQOTAFQVAI  
ENGWQETHOACRFHGVOHLEHLFYGANVAGNAGSNVAHHICALYNQESCARVLFLFRGANRDVRNYSQOTAFQVAI  
ENGWQETHOACRFHGVOHLEHLFYGHAFYGAONASGNVAHHICALYNQETCARHLFRGANRDVRNYSQOTAFQVAI

(B)

Human SHANK3  
Rat Shank3  
Rat Shank1A  
Rat Shank2

LOKRDHEGFGFVLRGAKAETPIEEFTPTPAFPALQYLESVDVEGVAMRAGLRTGDFLIEVGNVNVKVGHKQVVALIROGGNRLVMKVSVTRKPEEDSAR\_RRA  
LOKRDHEGFGFVLRGAKAETPIEEFTPTPAFPALQYLESVDVEGVAMRAGLRTGDFLIEVGNVNVKVGHKQVVALIROGGNRLVMKVSVTRKPEEDSAR\_RRA  
LOKRDSEGFGFVLRGAKAETPIEEFTPTPAFPALQYLESVDVEGVAMRAGLRTGDFLIEVGNVNVKVGHKQVVALIROGGNRLVMKVSVTRKPEEDSAR\_RRA  
LOKRDNEGFGFVLRGAKAETPIEEFTPTPAFPALQYLESVDVEGVAMRAGLRTGDFLIEVGNVNVKVGHKQVVALIROGGNRLVMKVSVTRKPEEDSAR\_RRA

(C)

Human SHANK3  
Rat Shank3  
Rat Shank1A  
Rat Shank2

WSKPEDVGMLESIIHIGEHRDRFEDHEIEGAHLPALTKDDFVELGVTRVYGHRRMNIERALKQLDGS  
WSKPEDVGMLESIIHIGEHRDRFEDHEIEGAHLPALTKDDFVELGVTRVYGHRRMNIERALKQLDGS  
MTKFDVADWLEMLCLSEHRAQFLDHEIDGSHLEALRKEDYDLGVTRVYGHRRMNIERALKPFLER  
MTKFDVADWLESLNLGEHKEFEMDNEIDGSHLENLQEDLIDLGVTRVYGHRRMNIERALKQLDDR

(D)

Human SHANK3  
Rat Shank3  
Rat Shank1A  
Rat Shank2

LVPPPEEFAN  
LVPPPEEFAN  
PLPPPEEFAN  
PLPPPEEFAN

(E)

Human SHANK3  
Rat Shank3  
Rat Shank1A

PGRKFTAVKAHSPQOGEETPLHRGEAVKVLISIGEGGFWEQTVKRTGTFWFPADCVBEVQVR  
PGRKFTAVKAHSPQOGEETPLHRGEAVKVLISIGEGGFWEQTVKRTGTFWFPADCVBEVQVR  
PGRSFWAKSVYQAQGEGETISLSKGEKIKVLISIGEGGFWEQTVKRTGTFWFPADCVBEVQVR

(F)

Human SHANK3  
Rat Shank3  
Rat Shank1A  
Rat Shank2

KPPVPPPKP  
KPPVPPPKP  
GTPPPPP  
KPPVPPPKP







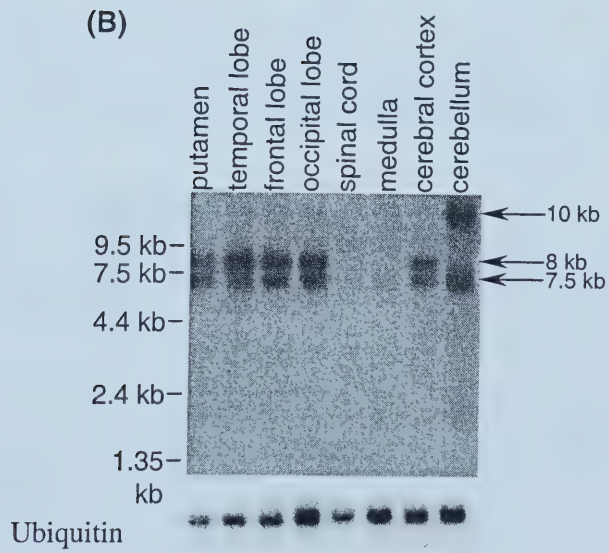
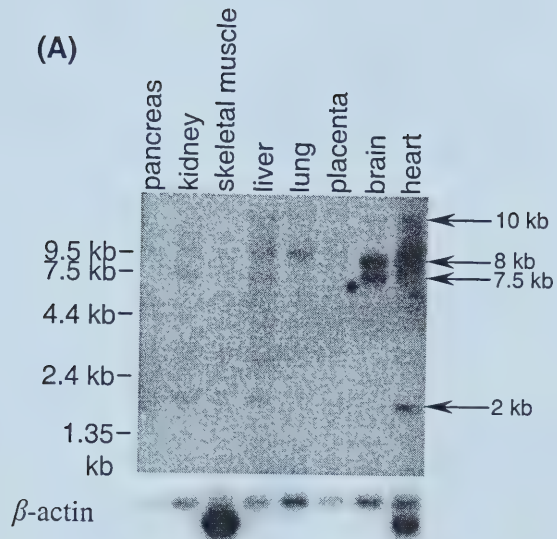
Figure 3-4 continued: Alignment of human SHANK3 domains with other Shank gene family members. Letters highlighted in black indicate identical amino acids, and letters highlighted in grey indicate conserved amino acids. (G) Proline Rich Domain. This domain is less conserved between the different Shank family members than the other protein interacting domains.







Figure 3-5: Northern Blot Analysis of SHANK3 in humans. Probe used contains exons 12 to 16 of SHANK3 gene. A. Probe on Clontech Multiple Tissue Northern Blot. A weak band of 2kb is observed in most of the lanes, as well as a large 10kb band showing a similar pattern of expression. An 8kb band is observed in most lanes tested with strongest expression in brain and heart. A 7.5kb band is observed strongly in brain and weaker in kidney. B. Probe on Clontech Brain Northern Blot. No 2kb band is observed on this blot. The 8kb band is observed in all lanes, but at very low levels in cerebellum, medulla and spinal cord. The 7.5kb band is observed in all lanes, but with weaker expression in the medulla and spinal cord. The large 10kb transcript is observed at low levels in all tissues, but shows high expression in the cerebellum.







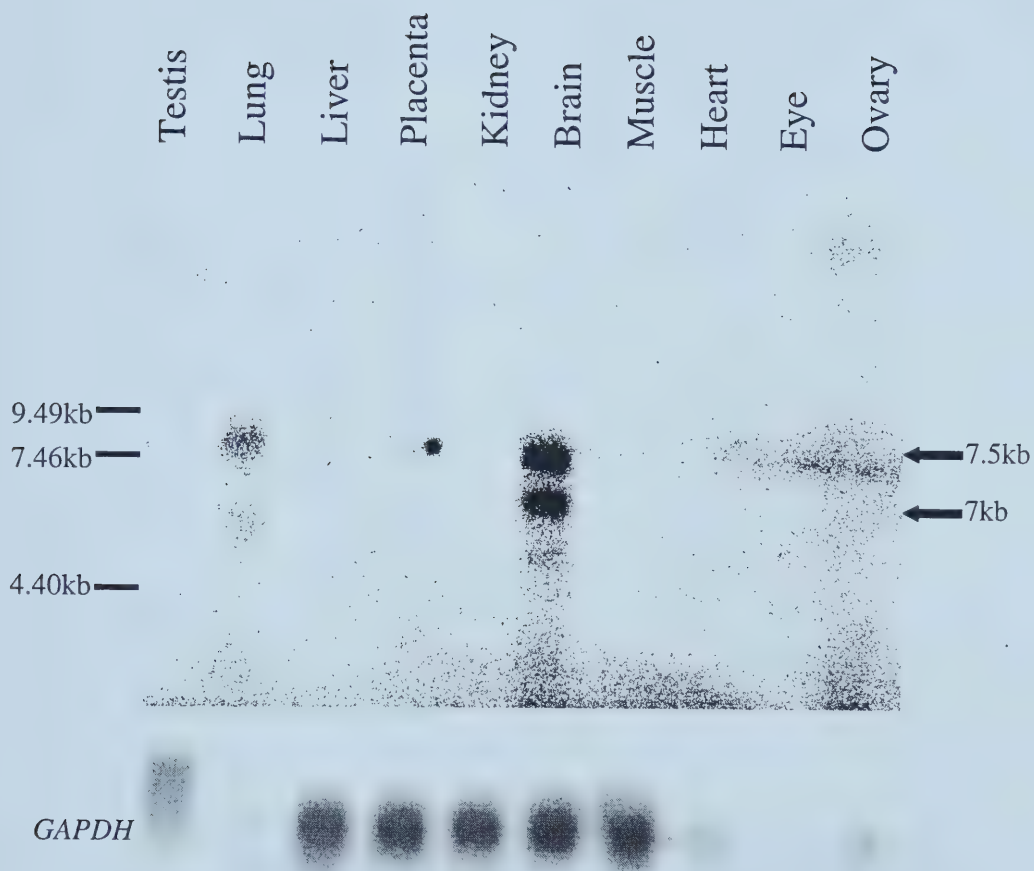


Figure 3-6: Mouse multiple tissue Northern blot. Northern blot probed with exons 12 to 16 of human *SHANK3*. Two bands as observed in the brain, one approximately 7kb and the other approximately 7.5kb. A very faint band of 7.5kb is visible in the lung.



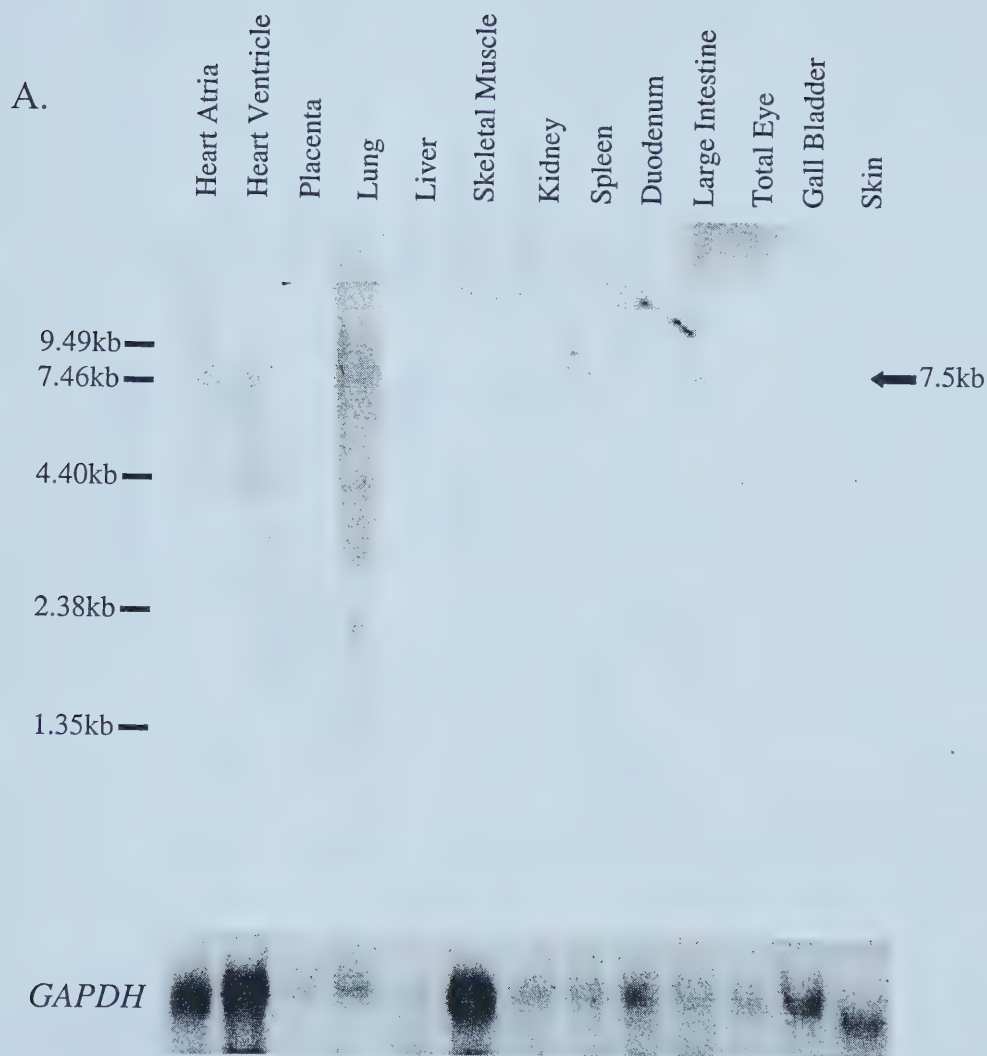


Figure 3-7: A. Pig multiple tissue Northern blot probed with exons 12 to 16 of human *SHANK3*. A band of 7.5kb is observed in the lung.



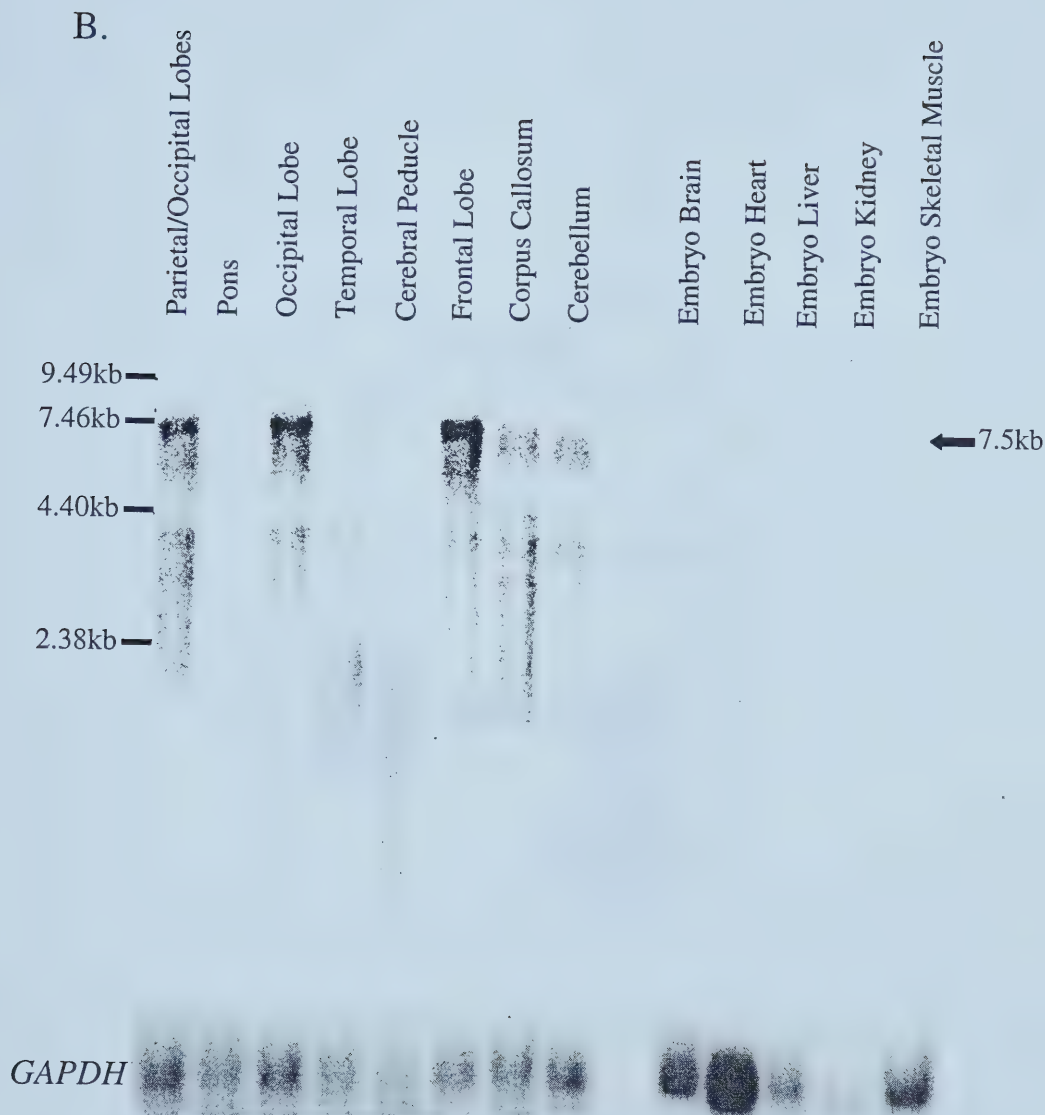


Figure 3-7 continued: B. Pig multiple tissue brain and embryo Northern blots probed with exons 12 to 16 of human *SHANK3*. A single band of 7.5kb is observed in several different regions of the brain including parietal/occipital lobe, occipital lobe, frontal lobe and possible expression in the temporal lobe, corpus callosum, cerebellum. On the embryo Northern blot, expression of the 7.5kb transcript is observed in brain, and possibly in heart and skeletal muscle.



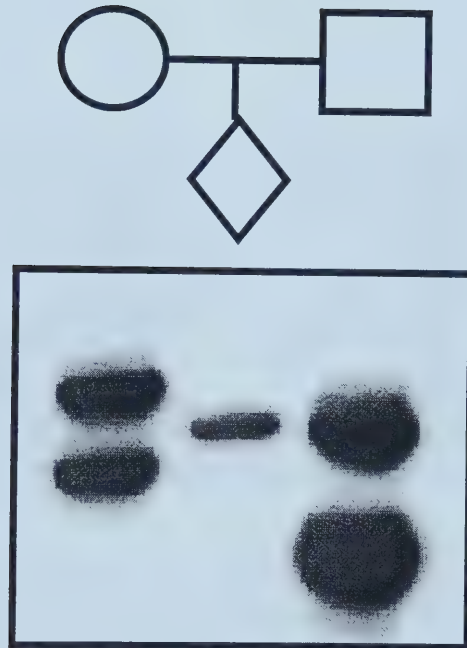


Figure 3-8: Example of a result for minisatellite D22S163. The DNA was digested with *Sau3A* and run on an agarose gel before Southern transfer to a membrane. Maternal DNA is in the left lane showing a heterozygote, paternal DNA is in the right lane showing a heterozygote and child (proband) DNA is in the center lane showing a single allele. This patient shows a maternal germline deletion of D22S163 probe.





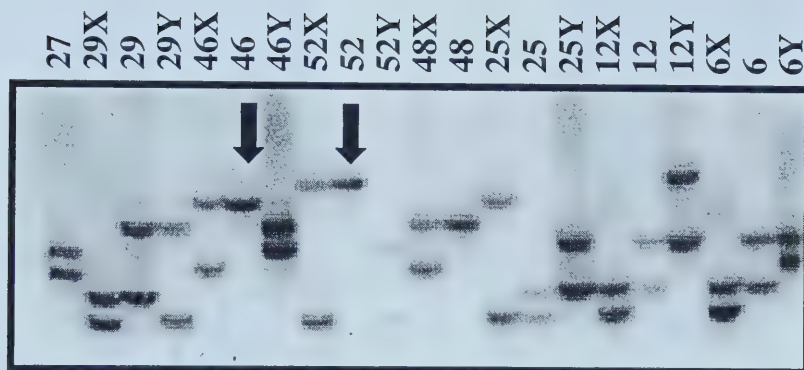
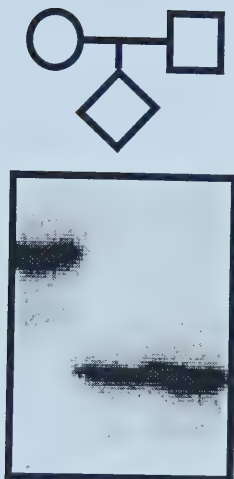


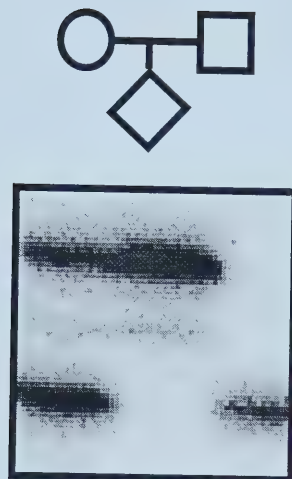
Figure 3-9: Example of a Microsatellite gel. Primers were endlabeled with radioactivity and a PCR was carried out on patient and parental DNA if available. PCR reactions were run on an 8% polyacrylamide gel to obtain good separation of bands. Each column represents an individual. Individuals are identified across the top by a number. The number represents the patient number and an X following the number indicates a maternal sample, and a Y following the number indicates a paternal sample. Both heterozygotes and hemizygotes (arrows) for this locus (D22S417) are observed. Patients 6, 12, 25, 27, and 29 are heterozygotes, indicating no deletion of this locus. Patient 46 is a hemizygote showing a maternal germline deletion. Patient 52 is a hemizygote showing a paternal germline deletion. Patient 48 does not have both parental DNA samples available therefore this patient is uninformative, it could be either a hemizygote or a homozygote.



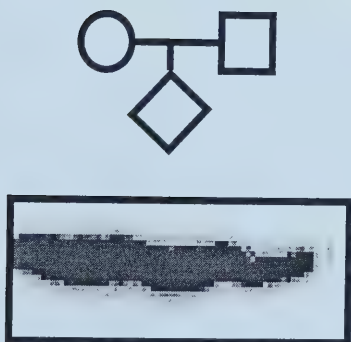
A.



B.



C.



D.

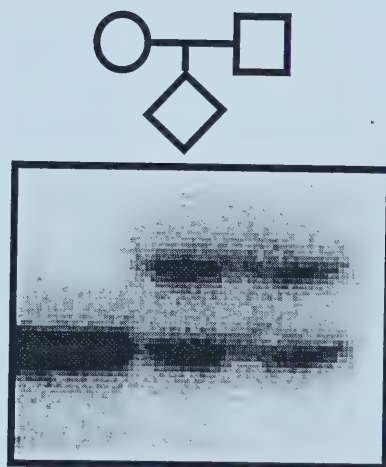


Figure 3-10: Examples of microsatellite analysis. A. Shows an example of a maternal germline deletion. B. Shows an example of a paternal germline deletion. C. Shows an example of an uninformative microsatellite. D. Shows an example of a microsatellite that is not deleted.





Figure 3-11: Example of FISH results. Arrows point to chromosomes with fluorescent signal. A. Only a single fluorescent signal is observed on this metaphase spread of 46 chromosomes. Therefore this probe, cN21F1, is deleted in patient 4. B. Two sets of fluorescent signal are observed on this metaphase spread. Therefore this probe, cN104C4, is not deleted in patient 4.

A.



B.









Figure 3-12: Summary of sizes of patient deletions. Each patient is represented by a horizontal line. The numbers down the right side are the numbers and abbreviations used to identify patients. A solid line indicates the region of a confirmed deletion. A dotted line indicates the uninformative region. A blue line represents a paternal germline deletion, a pink line a maternal germline deletion, and a green line the parent of origin of the deletion is unknown. The microsatellite markers are labeled across the top in purple. FISH cosmids used in this thesis are labeled across the top in teal, and other markers used are labeled in black. A wide range of sizes of deletions is observed in these patients, from 130kb to >10Mb. The MRS interstitial deletion is identified as the last patient and does not overlap with the smaller terminal deletions.



750q7



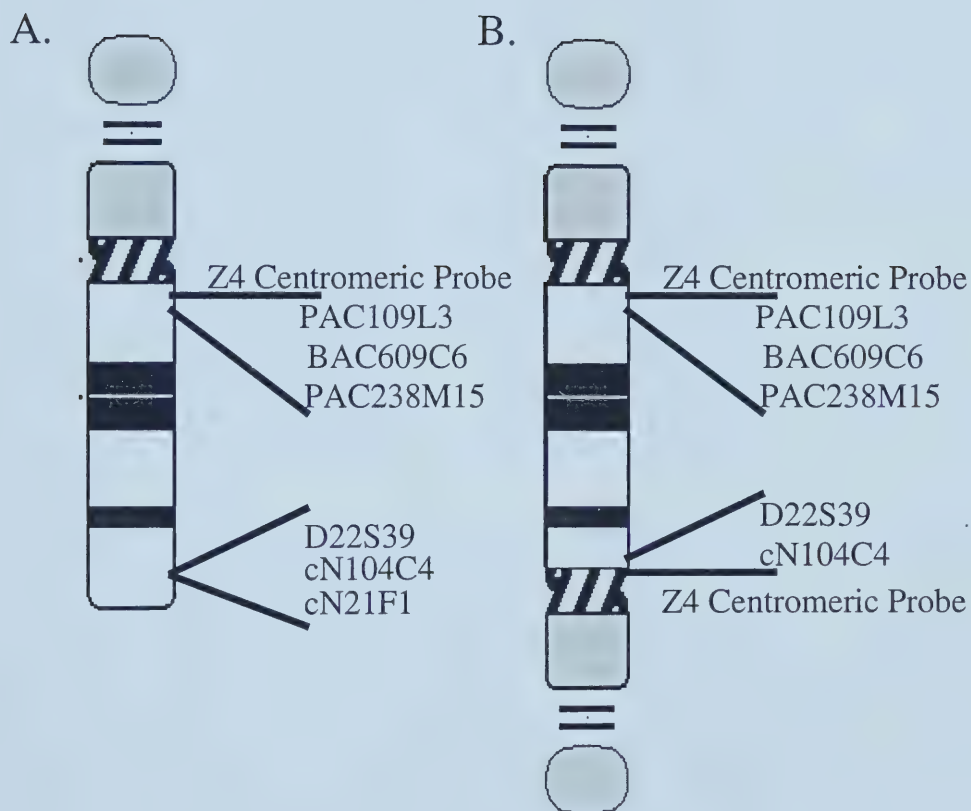


Figure 3-13: Patients 4's bisatellited chromosome. Comparison of bisatellited chromosome to a normal chromosome 22 with relative position of FISH probes used. A. Normal chromosome 22. B. Patient 4's bisatellited chromosome 22.

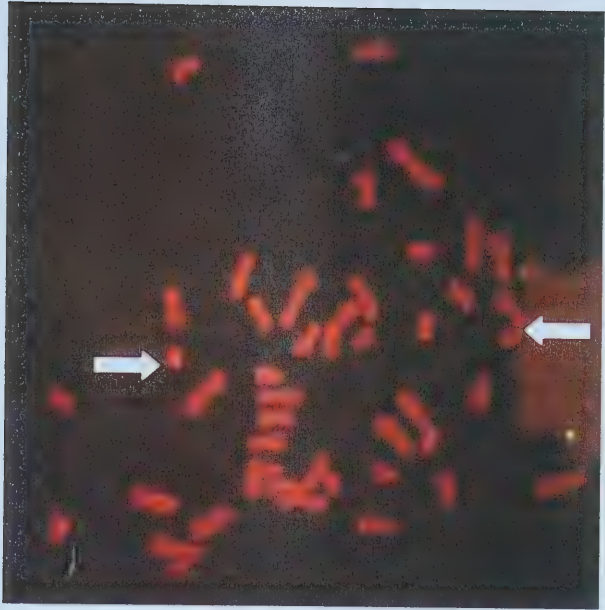






Figure 3-14: FISH results from patient 2. Signals are identified with arrows A: FISH with cosmid cN66C4. Two signals are observed in this metaphase preparation indicating that cN66C4 is not deleted in patient 2. The signal on right is not as strong as the signal on left, indicating a potential partial deletion of part of this cosmid. B: FISH with cosmid cN85A3. A single signal is observed on this metaphase preparation of patient 2 indicating a deletion of this cosmid.

A.



B.

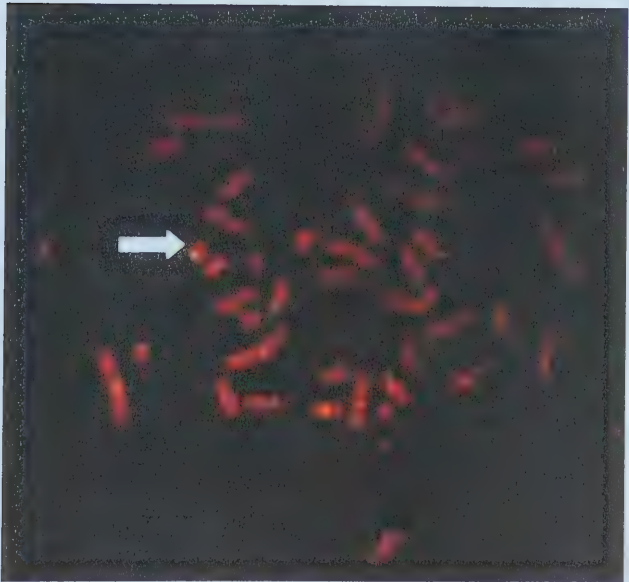




Table 3-1: PCR products obtained in genomic analysis of SHANK3

PCR Product Name	Forward Primer	Reverse Primer	Template	Exons Contained
start to sc24	start-1	sc24-3	Heart and Muscle	1, 2, 3, 4, 5
sc24	sc24-4	sc24-6	Heart and Muscle	2, 3, 4, break, 9, 10
sc24 to I511	sc24-2	I511-2	Heart and Muscle	8, 9, 10, 14, 15, 16
I511- 1st part	I511-12	I511-3	Heart and Muscle	12, 13, 14, 15, 16
I511	I511-5	I511-7	Heart and Muscle	12, 13, break, 19, 20, 21
Gen to FL2	gen1-2	FL2-1	Heart and Muscle	23, 25, 27
FLS	FLS-2	FLS-3	Heart and Muscle	24
FL2	FL2-2	FL2-3	Heart and Muscle	25, 27
<p><b>Legend:</b> <b>PCR Product Name:</b> name given to identify PCR product; <b>Forward Primer:</b> name given to forward primer used in PCR reaction; <b>Reverse Primer:</b> name given to reverse primer used in PCR reaction; <b>Template:</b> tissue source of RNA used in RT reaction; <b>Exons Contained:</b> number of the exon(s) sequenced that are contained within the PCR product; break means a gap in the sequence of the PCR product</p>				





Table 3-2: Summary of Minisatellite and Microsatellite Analysis

Number	D22S163	D22S526	D22S1169	D22S1170	D22S1149	D22S532	D22S1141	D22S928	D22S274	D22S1159	D22S1171	D22S927
1/N/T	I	2	2	2	U	U	2	U	2	U	2	U
2	ND	2	U	U	2	2	U	U	2	ND	2	2
3	I	2	2	I	ND	2	I	2	2	2	ND	I
4	I	IM	IM	U	U	U	2	2	2	U	2	U
5	IP	IP	U	U	U	2	U	2	2	U	2	ND
6	IP	I	IP	2	2	2	2	2	2	U	2	ND
7	I	I	I	I	2	I	I	I	I	I	2	I
8	U	IP	U	IP	2	2	U	2	U	U	ND	2
9	IM	IM	U	IM	2	2	2	U	2	2	ND	2
10/LM	I	I	I	ND	2	I	2	2	2	2	ND	ND
11	I	I	I	I	I	I	I	2	I	I	I	2
12	U	IM	IM	U	IM	2	2	2	2	2	U	ND
13	IM	U	U	U	IM	2	U	2	2	U	U	U
14	I	I	I	I	I	I	2	2	I	I	ND	2
15/CH	I	I	I	I	I	I	2	I	2	2	ND	2
16	U	U	IP	IP	U	U	IP	U	2	U	2	ND
17/AJ	U	IM	IM	IM	IM	IM	U	2	2	2	2	ND
18/JT	I	I	I	I	I	I	I	2	2	I	2	I
19	IP	IP	U	I	U	U	U	U	2	U	2	ND
20	IM	IM	IM	U	IM	IM	IM	IM	IM	U	2	2
21	IP	IP	U	IP	ND	IP	ND	IP	IP	2	ND	2
22/QM	I	I	I	I	I	I	I	I	I	ND	I	I
23	I	I	I	I	I	I	I	I	I	I	2	I
24	IM	IM	U	U	IM	U	IM	IM	IM	U	2	U
25	I	IP	IP	U	U	U	IP	U	IP	U	2	2
26	I	ND	ND	IP	ND	IP	IP	IP	I	I	ND	2
27	I	I	I	I	I	I	I	I	I	I	I	2
28	ND	ND	I	I	I	I	I	I	I	I	I	2
29	U	I	U	U	IP	U	IP	U	U	IP	IP	2
30	I	IP	U	U	U	IP	U	U	IP	U	IP	2
31	U	IP	IP	IP	U	IP	U	IP	U	U	IP	U
32	I	IP	IP	IP	I	U	U	U	U	ND	IP	U
33	I	U	IP	U	ND	IP	IP	IP	IP	U	ND	IP
34	I	I	I	I	I	I	I	I	I	I	I	I
35/CB	I	ND	IP	IP	IP	U	IP	IP	IP	I	IP	ND
36	U	U	IP	IP	ND	IP	ND	IP	U	ND	ND	U
37	IP	IP	IP	IP	IP	U	IP	IP	U	IP	IP	U
38	ND	ND	IM	U	ND	U	ND	I	IM	ND	ND	U
39/FB	I	IP	IP	IP	IP	U	U	IP	IP	IP	ND	IP
40	ND	I	I	I	ND	I	I	I	I	I	ND	I
41	IP	ND	IP	IP	U	U	ND	I	IP	IP	ND	IP
42	U	IM	U	IM	IM	U	ND	I	U	IM	ND	IM
43	U	IP	U	IP	U	U	IP	IP	IP	U	IP	IP
44	ND	I	I	I	ND	I	I	I	I	I	I	I
45	I	I	I	ND	ND	I	I	I	I	I	ND	I
46	IP	U	IP	IP	IP	IP	IP	U	U	U	IP	IP
47	U	IP	IP	U	ND	U	ND	ND	IP	IP	ND	U
48	I	I	I	I	I	I	I	I	I	I	I	I
49/NS	I	IM	IM	IM	IM	U	IM	U	U	I	IM	IM
50	IM	IM	IM	U	U	IM	IM	U	U	IM	IM	IM
51	IP	U	U	U	ND	IP	ND	IP	U	ND	ND	IP
52	U	I	IP	I	U	IP	IP	U	IP	I	IP	U
53	I	I	I	I	I	I	I	I	I	I	I	I
54	IP	IP	IP	IP	ND	U	ND	IP	IP	ND	ND	IP
55	U	IP	U	U	ND	IP	ND	I	IP	I	ND	IP
JxA	ND	2	2	2	U	U	2	2	2	ND	2	2
MRS	2	2	2	U	U	2	2	U	IM	ND	ND	IM
% Uninformative	34%	20%	34%	43%	50%	45%	27%	37%	29%	59%	8%	31%

**Legend:** Patients are identified in the first column by number or number/initials if previously published. Minisatellites and microsatellites are identified across the first row as D22S... 2 = heterozygote (2 alleles for locus); IM = maternal deletion (only one allele); IP = paternal deletion (only one allele); I = one allele observed, no parental information available (could be a homozygote or a hemizygote); U = locus is uninformative; ND = locus not done

Table 3-2: Summary of Minisatellite and Microsatellite Analysis

Number	D22S282	D22S1179	D22S417	D22S1166	D22S270	D22S1178	D22S1157	D22S276	D22S279	D22S423	D22S284	D22S272	D22S1177	D22S283
I/NT	2	2	2	ND	2	2	2	2	2	2	2	2	ND	2
2	ND	ND	U	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
3	2	ND	I	ND	2	ND	ND	ND	ND	ND	ND	ND	ND	ND
4	2	2	U	ND	2	U	2	2	2	2	2	2	2	2
5	2	2	2	ND	2	ND	ND	ND	ND	ND	ND	ND	ND	ND
6	2	U	2	ND	ND	2	ND	ND	ND	ND	ND	ND	ND	ND
7	I	I	I	ND	I	I	ND	2	2	I	2	2	ND	2
8	U	2	2	2	2	ND	ND	ND	ND	ND	ND	ND	ND	ND
9	2	ND	ND	ND	2	ND	ND	ND	ND	ND	ND	ND	ND	ND
10/LM	ND	ND	I	ND	I	ND	ND	ND	ND	ND	ND	ND	ND	ND
11	2	2	2	2	2	2	I	I	I	2	2	I	ND	2
12	2	U	2	ND	ND	ND	2	ND	ND	ND	ND	ND	ND	ND
13	2	2	2	2	ND	2	2	2	U	U	2	2	ND	2
14	ND	ND	2	I	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
15/CH	2	ND	I	ND	2	ND	ND	ND	ND	ND	ND	ND	ND	ND
16	2	2	2	ND	U	U	U	ND	ND	2	ND	ND	ND	ND
17/AJ	2	2	2	ND	2	2	U	ND	ND	2	ND	ND	ND	ND
18/JT	2	2	2	2	ND	2	I	2	I	I	2	2	ND	2
19	2	2	U	ND	U	U	U	ND	ND	ND	ND	ND	ND	ND
20	2	2	2	ND	2	2	2	ND	ND	2	ND	ND	ND	ND
21	U	ND	U	ND	2	ND	ND	ND	ND	ND	ND	ND	ND	ND
22/QM	2	2	2	ND	2	I	I	I	I	2	2	I	ND	2
23	2	2	I	ND	I	2	2	I	I	2	2	I	2	2
24	U	2	2	U	ND	ND	2	ND	ND	2	ND	ND	ND	ND
25	2	2	2	ND	ND	ND	2	ND	ND	2	ND	ND	ND	ND
26	2	2	I	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
27	I	2	2	I	2	2	ND	2	2	I	2	I	ND	2
28	2	I	2	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
29	U	2	2	ND	2	2	2	ND	ND	2	ND	ND	ND	ND
30	2	U	2	ND	2	ND	ND	ND	ND	ND	ND	ND	ND	ND
31	2	2	2	ND	2	ND	U	ND	ND	ND	ND	ND	ND	ND
32	2	2	2	ND	2	2	I	ND	ND	2	ND	ND	ND	ND
33	2	2	2	ND	2	ND	ND	ND	ND	ND	ND	ND	ND	ND
34	I	2	2	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
35/CB	IP	2	2	2	2	U	ND	U	ND	ND	ND	ND	ND	ND
36	U	2	2	U	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
37	IP	2	2	2	2	2	ND	2	U	2	2	2	ND	2
38	IM	2	2	2	ND	ND	ND	ND	ND	ND	2	ND	2	2
39/FB	U	U	2	2	2	2	ND	2	ND	2	ND	ND	ND	ND
40	I	I	2	2	2	ND	ND	ND	ND	ND	ND	ND	ND	ND
41	U	IP	2	2	2	ND	ND	ND	ND	ND	ND	ND	ND	ND
42	U	IM	2	ND	I	ND	ND	ND	ND	ND	ND	ND	ND	ND
43	IP	IP	2	2	2	2	2	ND	ND	2	ND	ND	ND	ND
44	I	I	I	2	ND	2	ND	ND	ND	ND	I	2	2	2
45	I	ND	I	2	2	I	ND	I	2	2	I	2	ND	2
46	U	IP	IP	2	2	2	ND	2	2	ND	2	2	ND	2
47	IP	ND	IP	2	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
48	I	I	I	ND	I	2	2	ND	ND	ND	ND	ND	ND	ND
49/NS	U	IM	IM	U	U	2	2	2	ND	ND	ND	ND	ND	ND
50	IM	U	IM	ND	IM	U	2	2	2	2	U	2	ND	U
51	IP	ND	IP	ND	IP	ND	ND	2	2	2	U	2	U	U
52	IP	U	IP	ND	IP	I	ND	2	2	2	2	I	ND	2
53	I	I	I	I	I	IP	I	2	I	2	2	2	2	2
54	IP	IP	U	IP	ND	ND	ND	ND	2	2	2	2	2	2
55	IP	IP	IP	IP	ND	ND	ND	ND	ND	ND	2	2	U	2
JaA	2	2	2	ND	2	2	ND	U	U	2	2	2	ND	2
MRS	IM	IM	IM	IM	IM	U	ND	2	2	2	2	2	2	ND
%Uninformative	24%	16%	15%	21%	10%	27%	31%	19%	38%	5%	25%	0%	25%	13%

**Legend:** Patients are identified in the first column by number or number/initials if previously published. Minisatellites and microsatellites are identified across the first row as D22S... . 2 = heterozygote (2 alleles for locus); IM = maternal deletion (only one allele); IP = paternal deletion (only one allele); I = one allele observed, no parental information available (could be a homozygote or a hemizygote); U = locus is uninformative; ND = locus not done







Table 3-3: Summary of FISH Analysis

Number	Metaphase Preparation?	cN1G3	c202	cN85A3	cN66C4	cN4F12	cN55B9	cN125P7	cN113P5	cN63E9	cN62A6	cN49E7	cN11F1	A39 (cN68C3/cN108A7)	cN104C4	cN4D6	cN76E8
1	Yes		1														
2	Yes	1		1	2												
3	Yes					1			2					2			
4	Yes										1	1	1	2	2	2	
5	Yes													1	1	2	
6	Yes													1	1	2	3
7	No																
8	Yes		1														
9	Yes																
10LM	No									1							
11	Yes				1												
12	Yes				1												
13	Yes																
14	Yes																
15CH	Yes				1												
16	No																
17(A)	Yes		1		1												
18/TT	Yes																
19	No																
20	Yes																
21	Yes																
22QM	Yes																
23	No																
24	Yes																
25	No																
26	Yes																
27	Yes				1												
28	Yes				1												
29	Yes																
30	No									1							
31	No																
32	Yes				1												
33	Yes																
34	No																
35CB	No									1							
36	Yes																
37	Yes																
38	Yes				1												
39FB	Yes																
40	Yes				1												
41	Yes																
42	Yes																
43	No																
44	Yes																
45	No																
46	Yes																
47	Yes	1															
48	No																
49/NS	No																
50	Yes																
51	No																
52	Yes				1												
53	Yes				1												
54	Yes																
55	No																
76A	Yes																
MBS	Yes																

Legend: Fluorescent *In Situ* Hybridization analysis of patient samples. Patients are listed in the first column by number or number/abbreviation if published. Cosmids used for FISH analysis are listed across the first row beginning with cN... The second column: Metaphase Preparations indicates if metaphase cell preparations were available on the patient or not. 1 = if a single set of fluorescent signal was observed on a metaphase spread, then the cosmid is deleted from one of the patient's chromosome 22's; 2 = if two fluorescent signals are observed on a metaphase spread, then the cosmid probe is not deleted in this patient.

Table 3-3: Summary of FISH Analysis

Number	cN6A04	cN7C5	cN63C4	cN119A3	cN75H12	cN65D1	cN79E2	cN104F9	cN24A8	cN31E2	cN9K2	cN128A12	cN75B3	cN10H1	cN31B6	cN61A1	cN17H6
18T							2										
2																	
3																	
4	2																
5																	
6																	
7																	
8				1		1		1									
9					1	1	1										
10LM																	
11								2	2								
12								2									
13								2									
14																	
15CH										1							
16																	
17AJ											1						
18/T																	
19		1															
20												2					
21												1	2				
22QM																	
23																	
24																	
25																	
26												1					
27												1					
28												1					
29																	
30																	
31																	
32																	
33																	
34																	
35CB																	
36																	
37																	
38														1			
39FB														1			
40													1	1			
41																	
42																	
43																	
44																	
45																	
46																	
47															2		
48																	
49VS																	
50																	1
51																	
52															1	1	1
53																1	1
54																1	
55																	
JaA																	
MR3																1	1

**Legend:** Fluorescent *In Situ* Hybridization analysis of patient samples. Patients are listed in the first column by number or number/abbreviation if published. Cosmids used for FISH analysis are listed across the first row beginning with cN... . The second column: Metaphase Preparations indicates if metaphase cell preparations were available on the patient or not. 1 = if a single set of fluorescent signal was observed on a metaphase spread, then the cosmid is deleted from one of the patient's chromosome 22's; 2 = if two fluorescent signals are observed on a metaphase spread, then the cosmid probe is not deleted in this patient.





Table 3-4: Summary of parent of origin, karyotype and molecular breakpoints of patients

Number	Parent of Origin	Karyotype	Distal Marker of Breakpoint	Proximal Marker of Breakpoint	Approximate Size of Deletion in kb	Size of Uninformative Region in kb
1/NT	PAT		cN66C4	cN66C4	130	
2	MAT	46, XX, der(22)t(14;22)(q32.31;q13.33)mat	cN66C4	cN66C4	130	unknown
3	Unknown		cN4F12	cN113F5	917	757
4	MAT		D22S1169	cN104C4	1972	221
5	PAT	46, XX, del(22)(q13.3) de novo	cN104C4	cN4D6	2323	351
6	PAT	46, XY, ish del(22)(q13.3q13.3)(D22S39-) de novo	cN104C4	cN4D6	2323	351
7	Unknown		Below D22S1149		4496	
8	PAT		cN65D1	D22S1149	4496	930
9	MAT	46, XX, der(22)t(19;22) mat	cN79E2	D22S1149	4496	639
10/LM	MAT		D22S294	D22S1149	4496	596
11	Unknown		Below cN104F8		4600	
12	MAT	46, XX, der(22)t(1;22)(q44;q13.32)	D22S1149	cN104F8	4600	104
13	MAT		cN104F8	D22S532	4984	384
14	Unknown	46, XX, del(22)(q13.3) de novo	Below D22S1141		5388	
15/CH	Unknown		cN38E2	D22S1141	5388	183
16	PAT	46, XX, del(22)(q13.3) de novo	D22S1141	D22S97	5550	162
17/AJ	MAT		cN89G2	D22S928	5632	132
18/JT	Unknown		D22S97	D22S928	5632	82
19	PAT	46, XX, del(22)(q13.31) de novo	D22S526	D22S274	5838	4524
20	MAT	46, XX, der(22)t(22;acro)(q13.33;p12) de novo	D22S274	cN128A12	6176	338
21	PAT	46, XY, del(22q13) de novo	cN128A12	D22S1159	6357	181
22/QM	Unknown	46, XY, del(22)(q13.31)	D22S97	D22S40	6442	892
23	Unknown		Below D22S1171		6708	
24	MAT	46, XY, del(22)(q13.1q13.3) de novo	D22S274	D22S1171	6708	870
25	PAT	46, XY, del(22)(q13.3) de novo	D22S40	D22S1171	6708	266
26	PAT	46, XX, der(22)t(22;acro).ish del(22)(q13.3) de novo	cN128A12	D22S927	7041	865
27	Unknown	46, XY, del(22)(q13.31)	cN128A12	D22S927	7041	865
28	Unknown	46, XY, del(22)(q13.3) de novo	cN128A12	D22S927	7041	865
29	PAT	46, XX, del(22)(q13.3) (D22S39-) de novo	D22S1171	D22S927	7041	333
30	PAT		D22S1171	D22S927	7041	333
31	PAT	46, XX, der(22)t(20;22)(p13.2;q13.3) pat	D22S1171	D22S282	7319	611
32	PAT	46, XY, der(22)t(22;acro)(q13.11;p14)[4]/47, XY, der(22)t(22;acro),+mar[16] de novo	D22S1171	D22S282	7319	611
33	PAT		D22S927	D22S282	7319	278
34	Unknown	46, XX, del(22)(q13.31) de novo	D22S40	D22S64	7400	958
35/CB	PAT	46, XX, del(22)(q13.3) de novo	D22S282	D22S64	7400	81
36	PAT		D22S928	D22S1179	7548	1916
37	PAT	46, XX, del(22)(q13.2) de novo	D22S282	D22S1179	7548	229
38	MAT	46, XX, del(22)(q13) de novo	cN10H1	D22S1179	7548	148

**Legend:** Number: is number or abbreviation used to identify patients; Parent of Origin: MAT = deletion occurred on maternal chromosome; PAT = deletion occurred on paternal chromosome; Unknown = parent of origin of deletion not available; Karyotype: karyotype written using conventional cytogenetic terminology; Distal Marker of Breakpoint: the most proximal marker that shows a deletion; Proximal Marker of Breakpoint: the most distal marker that is heterozygous; Approximate Size of Deletion in kb: the maximum size of the deletion measured from proximal marker of breakpoint to subtelomeric repeats; Size of Uninformative Region in kb: the size of the region between the proximal marker of breakpoint and distal marker of breakpoint

Table 3-4: Summary of parent of origin, karyotype and molecular breakpoints of patients

Number	Parent of Origin	Karyotype	Distal Marker of Breakpoint	Proximal Marker of Breakpoint	Approximate Size of Deletion in kb	Size of Uninformative Region in kb
39/FB	PAT	46, XY, del(22)(q13.3) de novo	D22S927	BZRP	7582	263
40	Unknown	46, XY, del(22)(q13.2) de novo	cN10H1	D22S417	8037	637
41	PAT	46, XX, del(22)(q13.3) de novo	D22S1179	D22S417	8037	489
42	MAT	46, XY, del(22)(q13.3) de novo	D22S1179	D22S417	8037	489
43	PAT		D22S95	D22S417	8037	287
44	Unknown	46, XY, del(22)(q13.2) de novo	Below D22S1166		8087	
45	Unknown		D22S95	D22S1166	8087	337
46	PAT	46, XX, del(22)(q13.3) de novo	D22S417	D22S1166	8087	80
47	PAT	46, XX, del(22)(q13.3) de novo	D22S417	D22S1166	8087	50
48	Unknown	46, XX, del(22)(q13.31) de novo	Below D22S1178		8738	
49/NS	MAT	46, XY, del(22)(q13.3) de novo	DIA1	D22S1178	8738	642
50	MAT	46, XY, del(22)(q13.3) de novo	cN17H6	D22S1157	8889	289
51	PAT		D22S270	D22S276	9116	1028
52	PAT	46, XYqs, del(22)(q13.3) de novo	cN17H6	D22S276	9116	516
53	PAT	46, XY, del(22)(q13.3) de novo	D22S1178	D22S276	9116	378
54	PAT	46, XX, del(22)(q13) de novo	cN61A1	D22S279	10, 106	1706
55	PAT	46, XX, del(22)(q13.31) de novo	D22S1166	D22S284	10, 812	2725
JaA		46, XY, del(22)(q13)[30%]/46, XY [70%]				
MRS proximal	MAT		cN17H6	D22S276	9116	516
MRS distal	MAT		D22S1141	D22S274	5838	449

**Legend:** **Number:** is number or abbreviation used to identify patients; **Parent of Origin:** MAT = deletion occurred on maternal chromosome; PAT = deletion occurred on paternal chromosome; Unknown = parent of origin of deletion not available; **Karyotype:** karyotype written using conventional cytogenetic terminology; **Distal Marker of Breakpoint:** the most proximal marker that shows a deletion; **Proximal Marker of Breakpoint** the most distal marker that is heterozygous; **Approximate Size of Deletion in kb:** the maximum size of the deletion measured from proximal marker of breakpoint to subtelomeric repeats; **Size of Uninformative Region in kb:** the size of the region between the proximal marker of breakpoint and distal marker of breakpoint





7

.

Table 3-5: Summary of clinical features of deletion patients

Number	Sex	Parent of Origin	Age in months	SIB BI AE	SIB BI AE	SIB MS AE	SIB PL AE	SIB CL AE	SIB SS	DP P	DP SH	DP S	DP A	DP C	DP IQ	Words Spoken	Words Recognized	Hypotonia	Ability to Walk	Growth	Head circumference	Organ Malformations	Neurological Scan	Seizures	
1/NT	M	PAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	+	-	ND	ND	ND	ND	ND	ND	
2	F	MAT	84	NA	7	5	8	7	9	26	16	8	12	15	12	14	-	+	+	---	---	None	ND	-	
3	F	Unknown	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	---	ND	-	++	+	---	None	ND	-	
4	F	MAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	ND	ND	+	---	ND	ND	ND	
5	F	PAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	ND	ND	+	---	ND	ND	ND	
6	M	PAT	120	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	++	---	None	-	+	
7	Unknown	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	-	ND	ND	ND	ND	ND	ND	
8	F	PAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	+	ND	ND	ND	ND	-	ND	
9	F	MAT	72	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	---	---	None	ND	-	
10/LM	F	MAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	-	ND	ND	++	ND	ND	ND	ND	ND	
11	F	Unknown	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	-	+	++	+	+	ND	ND	-	
12	F	MAT	84	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	+	ND	ND	None	ND	-	
13	F	MAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	---	+	-	ND	---	---	ND	-	+	
14	F	Unknown	24	76	18	21	19	15	18	43	21	23	16	18	12	72	+	+	++	+++	+++	Kidney	ND	-	
15/CH	Unknown	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	---	ND	-	++	ND	+++	ND	ND	ND	
16	F	PAT	72	NA	9	7	10	12	10	29	12	20	20	21	14	22	+	ND	+	+	+	None	-	+	
17/AJ	MAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	---	ND	+	+	+	+++	ND	-	ND	
18/JT	Unknown	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	+	+	++	+	ND	+	ND	
19	F	PAT	48	40	21	22	16	26	19	35	24	26	20	24	22	48	ND	+	ND	++	+	None	ND	-	
20	F	MAT	72	NA	23	32	18	30	10	10	22	22	34	28	30	38	+++	ND	+	ND	++	++	None	ND	+
21	M	PAT	60	NA	23	32	18	30	10	10	22	22	34	28	30	38	+	ND	+	+	+	+	None	ND	-
22/QM	M	Unknown	192	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	---	ND	+	+	+	---	None	ND	+	
23	M	Unknown	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	-	ND	ND	ND	ND	ND	Brain	ND	ND	
24	M	MAT	108	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	+	ND	ND	None	ND	-	
25	M	PAT	ND	NA	18	17	12	15	9	25	26	18	14	15	14	<17	-	+	+	++	+++	None	-	-	
26	F	PAT	48	31	19	21	17	20	14	46	24	24	26	28	16	48	---	-	+	+	+	None	ND	-	
27	M	Unknown	120	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	---	+	None	ND	+	
28	M	Unknown	72	NA	15	22	12	15	6	39	26	32	30	26	14	35	---	ND	+	ND	+	None	ND	-	
29	F	PAT	60	17	18	28	16	18	9	32	ND	ND	ND	ND	ND	+	ND	+	ND	+	+	None	-	-	
30	PAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	

**Legend:** Patient: Number of code used to identify individuals; Sex: M = male; F = female; Parent of Origin: parent of origin of deletion; MAT = maternal deletion; PAT = paternal deletion; Unknown = unknown deletion; Age in Months: approximate age in months; Scales of Independent Behavior (SIB): SIB BI: Broad Independence score; NA = too low to score; SIB BI AE: Broad Independence score in age equivalent (months); SIB MS AE: Motor skills in months; SIB SC AE: Social/Communication skills in months; SIB PL AE: Personal Living skills in months; SIB CL AE: Community Living skills in months; SIB SS: Support score; Developmental Profile II (DP): DP P: Physical skills in months; DP SH: Self help skills in months; DP S: Social skills in months; DP A: Academic skills in months; DP C: Communication skills in months; DP IQ: Estimated Intelligence score from combined DP scales; Words Spoken: --- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Words Spoken: --- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Hypotonia: - = no hypotonia; Ability to walk: --- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Growth: --- = undergrowth; - = below average; + = average; ++ = above average; +++ = overgrowth; Head Circumference: --- = microcephaly; - = below average; + = average; ++ = above average; +++ = macrocephaly; Organ Malformations: organ affected by malformation listed; Neurological Scan: + = normal scan; - = abnormal scan; Seizures: + = seizures; - = no seizures; Hearing Defects: + = hearing defect; - = normal hearing; Hands and Feet: S = syndactyly; C = clinodactyly; TN = abnormal tonus; LH = large hands; Recurrent Ear Infections: + = recurrent ear infections; - = no recurrent ear infections; Cesarean Section: + = born by cesarean; - = normal birth; Respiratory Problems at Birth: + = respiratory problems at birth; - = no respiratory problems at birth; Eyes: EF = epicanthal folds; P = ptosis; DS = deep set eyes; FE = full eye lids; LL = long eye lashes; A = abnormal, but not specified; Ears: D = dysmorphic ears; L = long ears; LS = low set ears; S = simple ears; T = preauricular pits and/or tags; P = prominent; A = abnormal, but not specified; High Palate: + = highly arched palate; - = normal arch in palate; Pointed Chin: + = pointed chin; - = no pointed chin; Dental Anomalies: + = dental anomalies; - = no dental anomalies; Full Cheeks: + = full cheeks; - = no full cheeks; Bulbous Nose: + = bulbous nose; - = no bulbous nose; Wide Nasal Bridge: + = wide nasal bridge; - = no wide nasal bridge; Long Philtrum: + = long philtrum; - = no long philtrum; Dolicocephaly: + = dolicocephaly; - = no dolicocephaly; Recurrent Upper Respiratory Problems: + = recurrent upper respiratory problems; - = no recurrent upper respiratory problems

Table 3-5: Summary of clinical features of deletion patients

Number	Sex	Parent of Origin	Age in months	SIB BI	SIB BI AE	SIB SC AE	SIB PL AE	SIB CL AE	SIB SS	DP P	DP SH	DP S	DP A	DP C	DP IQ	Words Spoken	Words Recognized	Hypotonia	Ability to Walk	Growth	Head circumference	Organ Malformations	Neurological Scan	Seizures	
31	F	PAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	ND	ND	Heart	-	-	
32	M	PAT	96	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	-	++	++	Kidney	-	-	
33	M	PAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	Heart	-	-	
34	F	Unknown	132	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	+	+	Kidney	-	-	
35/CB	F	PAT	156	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	+	+	+++	+++	None	-	-	
36	M	PAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	-	ND	+	ND	ND	ND	Heart	ND	ND	
37	F	PAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	-	ND	+	ND	ND	ND	ND	+	
38	F	MAT	48													++	+	+	++	+	+	None	ND	-	
39/FB	M	PAT	ND	42	23	26	20	27	18	44	26	30	36	30	28	57	ND	ND	+	ND	ND	None	ND	-	
40	M	Unknown	72	NA	7	5	6	11	9	24	14	22	18	15	16	23	+	ND	+	-	+	+	None	ND	-
41	F	PAT	132	NA	16	21	14	18	10	26	ND	ND	ND	ND	ND	ND	ND	+	+	-	ND	None	ND	-	
42	M	MAT	84	NA	10	7	13	13	13	27	12	18	16	15	18	17	+	ND	+	ND	+	+	None	ND	-
43	M	PAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	-	+	-	ND	+	++	ND	ND	+	
44	M	Unknown	96	NA	8	8	6	11	9	22	14	8	14	15	12	12	-	+	+	-	-	None	ND	+	
45	F	Unknown	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	Heart	ND	ND	
46	F	PAT	120	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	+	ND	++	+	Heart	-	-	
47	F	PAT	48	NA	9	10	4	13	9	13	22	16	16	18	14	35	-	ND	+	ND	+	+	None	ND	-
48	F	Unknown	72	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	+	ND	ND	Kidney	-	+	
49/NS	M	MAT	132													-	+	+	+	+	++	Heart	-	+	
50	M	MAT	24	NA	16	23	16	12	10	28	22	16	32	23	18	16	-	+	+	ND	+	+	Kidney	ND	-
51	M	PAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	+	+++	ND	-	+	
52	M	PAT	72	NA	8	7	9	7	9	31	14	12	8	18	12	17	-	ND	+	+	-	-	None	ND	-
53	M	PAT	48	NA	9	9	8	11	10	27	14	14	10	18	6	28	ND	ND	+	ND	+	+	Kidney	-	-
54	F	PAT	12	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+	ND	+	+	+	None	ND	+
55	F	PAT	24	21	7	5	5	11	9	34	12	14	10	15	8	56	ND	ND	+	ND	+++	+	None	ND	-
JaA	M		84													ND	ND	+	ND	ND	ND	None	ND	-	
				NA	18	26	19	15	9	32	24	10	14	18	18	23	ND	ND	-	ND	ND	ND	-	-	
MRS	F	MAT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	-	ND	+	ND	ND	-	-	
Legend: Patient: Number of code used to identify individuals; Sex: M = male; F = female; Parent of Origin: parent of origin of deletion; MAT = maternal deletion; PAT = paternal deletion; Unknown = unknown deletion; Age in Months: approximate age in months; Scales of Independent Behavior (SIB): SIB BI: Broad Independence score; NA = too low to score; SIB BI AE: Broad Independence score in age equivalent (months); SIB MS AE: Motor skills in months; SIB SC AE: Social/Communication skills in months; SIB PL AE: Personal Living skills in months; SIB CL AE: Community Living skills in months; SIB SS: Support score; Developmental Profile II (DP): DP P: Physical skills in months; DP SH: Self help skills in months; DP S: Social skills in months; DP A: Academic skills in months; DP C: Communication skills in months; DP IQ: Estimated Intelligence score from combined DP scales; Words Spoken: -- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Hypotonia: - = hypotonia; + = no hypotonia; Ability to walk: -- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Growth: - = undergrowth; - = below average; + = average; ++ = above average; +++ = overgrowth; Head Circumference: -- = microcephaly; - = below average; + = average; ++ = above average; +++ = macrocephaly; Organ Malformations: organ affected by malformation listed; Neurological Scan: + = normal scan; - = abnormal scan; Seizures: + = seizure; - = no seizures; Hearing Defects: + = hearing defect; - = normal hearing; Hands and Feet S = syndactyly; C = clinodactyly; TN = abnormal toenails; LH = large hands; Recurrent Ear Infections: + = recurrent ear infections; - = no recurrent ear infections; Cesarean Section: + = born by cesarian; - = normal birth; Respiratory Problems at Birth: + = respiratory problems at birth; - = no respiratory problems at birth; Eyes/Eye = epicanthal folds; P = ptosis; DS = deep set eyes; FE = full eye lashes; LL = long eye lashes; A = abnormal, but not specified; Ears: D = dysmorphic ears; L = long ears; LS = low set ears; S = simple ears; T = preauricular pits and/tags; P = prominent; A = abnormal, but not specified; High Palate: + = highly arched palate; - = normal arch in palate; Pointed Chin: + = pointed chin; - = no pointed chin; Dental Anomalies: + = dental anomalies; - = no dental anomalies; Full Cheeks: + = full cheeks; - = no full cheeks; Bulbous Nose: + = bulbous nose; - = no bulbous nose; Wide Nasal Bridge: + = wide nasal bridge; - = no nasal nasal bridge; Long Philtrum: + = long philtrum; - = no long philtrum; Dolicocephaly: + = dolicocephaly; - = no dolicocephaly; Recurrent Upper Respiratory Problems: + = recurrent upper respiratory problems; - = no recurrent upper respiratory problems																									

**Legend:** Patient: Number of code used to identify individuals; Sex: M = male; F = female; Parent of Origin: parent of origin of deletion; MAT = maternal deletion; PAT = paternal deletion; Unknown = unknown deletion; Age in Months: approximate age in months; Scales of Independent Behavior (SIB): SIB BI: Broad Independence score; NA = too low to score; SIB MS AE: Motor skills in months; SIB SC AE: Social/Communication skills in months; SIB PL AE: Personal Living skills in months; SIB CL AE: Community Living skills in months; SIB SS: Support score; Developmental Profile II (DP): DP P: Physical skills in months; DP SH: Self help skills in months; DP S: Social skills in months; DP A: Academic skills in months; DP C: Communication skills in months; DP IQ: Estimated intelligence score from combined DP scales; Words Spoken: --- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Words Recognized: --- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Hypotonia: + = hypotonia; - = no hypotonia; Ability to walk: --- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Growth: --- = undergrowth; - = below average; + = average; ++ = above average; +++ = overgrowth; Head Circumference: --- = microcephaly; - = below average; + = average; ++ = above average; +++ = macrocephaly; Organ Malformations: organ affected by malformation listed; Neurological Scan: + = normal scan; - = abnormal scan; Seizures: + = seizures; - = no seizures; Hearing Defects: + = hearing defect; - = normal hearing; Hands and Feet: S = syndactyly; C = clinodactyly; TN = abnormal toenails; LH = large hands; Recurrent Ear Infections: + = recurrent ear infections; - = no recurrent ear infections; Cesarean Section: + = born by cesarean; - = normal birth; Respiratory Problems at Birth: + = respiratory problems at birth; - = no respiratory problems at birth; Eyes: EF = epicanthal folds; P = ptosis; DS = deep set eyes; FE = full eye lids; LL = long eye lashes; A = abnormal, but not specified; Ears: D = dysmorphic ears; L = long ears; LS = low set ears; S = simple ears; T = preauricular pits and/or tags; P = prominent; A = abnormal, but not specified; High Palate: + = highly arched palate; - = normal arch in palate; Pointed Chin: + = pointed chin; - = no pointed chin; Dental Anomalies: + = dental anomalies; - = no dental anomalies; Full Cheeks: + = full cheeks; - = no full cheeks; Bulbous Nose: + = bulbous nose; - = no bulbous nose; Wide Nasal Bridge: + = wide nasal bridge; - = no wide nasal bridge; Long Philtrum: + = long philtrum; - = no long philtrum; Dolicocephaly: + = dolicocephaly; - = no dolicocephaly; Recurrent Upper Respiratory Problems: + = recurrent upper respiratory problems; - = no recurrent upper respiratory problems





Table 3-5: Summary of clinical features of deletion patients

Number	Hearing Defects	Chewing behavior	High Pain Tolerance	Hands and Feet	Recurrent Ear Infections	Cesarean Section	Respiratory Problems at Birth	Eyes	Ears	High Palate	Pointed Chin	Dental Anomalies	Full Cheeks	Bulbous Nose	Wide Nasal Bridge	Long Philtrum	Dolicocephaly	Recurrent Upper Respiratory Problems
1/NT	-	ND	ND	ND	ND	+	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
2	-	+	-	TN, C, S	+	ND	ND	P, LL	D	-	+	-	-	+	-	-	-	-
3	-	ND	ND	ND	+	ND	+	EF, LL	Normal	ND	ND	+	ND	ND	-	ND	ND	+
4/IE	-	ND	ND	TN, C	ND	ND	ND	DS	PP	+	ND	+	ND	ND	+	ND	ND	ND
5	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
6	-	ND	+	Normal	ND	ND	ND	P, LL	D	+	+	-	-	+	-	-	-	-
7	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
8	ND	ND	ND	TN	ND	ND	+	ND	PP	ND	ND	+	ND	ND	ND	ND	ND	+
9	-	+	+	TN, C, S	ND	ND	ND	P, EF, LL	Normal	-	+	-	+	+	+	-	-	-
10/LM	ND	ND	ND	C	S	ND	ND	ND	S	ND	-	+	ND	ND	ND	ND	ND	ND
11	ND	ND	ND	ND	+	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
12	-	+	+	Normal	ND	ND	ND	EF	Normal	+	-	-	-	+	-	-	-	-
13	ND	ND	ND	TN, S	ND	ND	+	EF	Normal	-	ND	ND	ND	ND	ND	+	ND	+
14	-	ND	+	FH, TN	+	ND	+	P, EF	Normal	-	+	-	-	-	-	-	-	+
15/CH	ND	ND	ND	C	ND	ND	ND	P, EF	A	+	ND	ND	ND	ND	+	ND	ND	ND
16	+	+	+	FH, TN, S	ND	+	ND	P, FE	D	-	+	+	+	+	+	-	-	-
17/AJ	ND	ND	ND	ND	ND	+	ND	ND	A	ND	ND	ND	ND	ND	ND	ND	ND	ND
18/IT	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
19	-	+	-	FH, LH, TN	+	ND	ND	FE, DS	D	-	+	-	+	+	-	-	+	-
20	-	ND	+	FH, LH, TN, S	+	ND	ND	EF	D	-	-	-	+	+	-	-	+	-
21	-	ND	+	FH, LH, TN, S	ND	+	ND	EF	D	-	+	-	ND	-	-	-	-	+
22/QM	-	ND	+	FH, TN, C	+	+	ND	P, EF, FE	D	-	+	-	+	+	-	-	-	-
23	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
24	-	+	+	PH	ND	ND	ND	P, FE	ND	+	-	-	+	-	-	-	+	-
25	-	ND	+	FH, TN	ND	ND	+	P, EF, DS	D	+	+	+	+	-	-	+	+	+
26	-	ND	+	C	ND	+	ND	P, EF	Normal	-	-	-	+	+	-	-	-	-
27	-	+	+	FH, TN, S	ND	ND	ND	Normal	D	-	+	-	+	-	-	-	+	-
28	-	+	+	FH, LH, TN	ND	ND	ND	LL	D	-	+	-	+	+	+	+	+	-
29	-	+	+	TN	+	ND	ND	P, EF	Normal	+	+	-	-	+	-	-	+	-
30	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND

**Legend:** Patient: Number of code used to identify individuals; Sex: M = male; F = female; Parent of Origin: parent of origin of deletion; MAT = maternal deletion; PAT = paternal deletion; Unknown = unknown deletion; Age in Months: approximate age in months; Scales of Independent Behavior (SIB): SIB BI: Broad Independence score; NA = too low to score; SIB B1 AE: Broad Independence score in age equivalent (months); SIB MS AE: Motor skills in months; SIB SC AE: Social/Communication skills in months; SIB PL AE: Personal Living skills in months; SIB CL AE: Community Living skills in months; SIB SS: Support score; Developmental Profile II (DP): DP P: Physical skills in months; DP SH: Self help skills in months; DP S: Social skills in months; DP A: Academic skills in months; DP C: Communication skills in months; DP IQ: Estimated Intelligence score from combined DP scales; Words Spoken: --- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Words Spoken: --- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Hypotonia: + = hypotonia; - = no hypotonia; Ability to walk: --- = very poor; - = poor; + = fair; ++ = good; +++ = very good; Growth: --- = undergrowth; - = below average; + = average; ++ = above average; +++ = overgrowth; Head Circumference: --- = microcephaly; - = below average; + = average; ++ = above average; +++ = macrocephaly; Organ Malformations: organ affected by malformation listed; Neurological Scan: + = normal scan; - = abnormal scan; Seizures: + = seizures; - = no seizures; Hearing Defects: + = hearing defect; - = normal hearing; Hands and Feet: S = syndactyly; C = clinodactyly; TN = abnormal toenails; LH = large hands; Recurrent Ear Infections: + = recurrent ear infections; - = no recurrent ear infections; Cesarean Section: + = born by cesarian; - = normal birth; Respiratory Problems at Birth: + = respiratory problems at birth; - = no respiratory problems at birth; Eyes: EF = epicanthal folds; P = ptosis; DS = deep set eyes; FE = full eye lids; LL = long eye lashes; A = abnormal, but not specified; Ears: D = dysmorphic ears; L = long ears; LS = low set ears; S = simple ears; T = preauricular pits and/or tags; P = prominent; A = abnormal, but not specified; High Palate: + = highly arched palate; - = normal arch in palate; Pointed Chin: + = pointed chin; - = no pointed chin; Dental Anomalies: + = dental anomalies; - = no dental anomalies; Full Cheeks: + = full cheeks; - = no full cheeks; Bulbous Nose: + = bulbous nose; - = no bulbous nose; Wide Nasal Bridge: + = wide nasal bridge; - = no wide nasal bridge; Long Philtrum: + = long philtrum; - = no long philtrum; Dolicocephaly: + = dolicocephaly; - = no dolicocephaly; Recurrent Upper Respiratory Problems: + = recurrent upper respiratory problems; - = no recurrent upper respiratory problems

Table 3-5: Summary of clinical features of deletion patients

Number	Hearing Defects	Chewing behavior	High Pain Tolerance	Hands and Feet	Ear Infections	Cesarian Section	Respiratory Problems at Birth	Eyes	Ears	High Palate	Pointed Chin	Dental Anomalies	Full Cheeks	Bulbous Nose	Wide Nasal Bridge	Long Philtrum	Dolicocephaly	Recurrent Upper Respiratory Problems
31	-	+	+	FH, S	ND	ND	ND	P, LL	D	-	+	-	+	-	-	-	+	-
32	-	ND	-	FH, TN, S	+	ND	+	P, EF, FE	D, L	+	+	-	+	-	-	-	+	+
33	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
34	-	ND	+	FH, TN, S	ND	ND	ND	P, FE	D	-	+	+	+	-	-	-	-	-
35/CB	-	ND	+	FH, TN, S	ND	+	ND	P, EF, FE, DS	D	+	+	-	+	+	-	+	+	-
36	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	+
37	-	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
38	-	+	+	TN	ND	+	ND	P, EF, FE	Normal	-	+	-	+	+	-	-	+	-
39/FB	-	+	+	FH, TN, S	ND	ND	ND	P, EF	D	-	+	-	-	+	+	-	+	-
40	-	+	+	FH, TN, S	ND	+	ND	P, EF	D	-	+	-	-	-	-	-	-	+
41	-	+	+	TN	ND	ND	ND	P, FE	D	-	-	-	-	-	+	-	-	-
42	-	+	+	TN, S	ND	ND	ND	P, EF	D	-	-	-	-	+	-	-	-	-
43	ND	ND	ND	S	ND	+	+	ND	D	ND	ND	+	ND	ND	ND	ND	+	ND
44	-	+	+	FH, TN, C	+	ND	ND	EF, LL	Normal	-	-	-	-	-	-	-	+	-
45	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
46	-	ND	+	Normal	ND	ND	ND	P	D	-	-	-	+	-	-	-	-	+
47	-	+	+	TN	+	ND	ND	P	Normal	-	-	-	-	-	-	-	-	-
48	-	ND	+	FH, TN	ND	ND	ND	P, EF	Normal	-	-	-	+	-	-	-	+	-
49/NS	-	+	+	FH, TN, C, S	ND	ND	+	P, LL, DS	D, PP	+	+	+	+	+	-	-	+	-
50	+	+	+	TN	ND	ND	ND	EF, LL	D	-	-	-	+	+	-	+	-	-
51	ND	ND	ND	ND	+	+	+	ND	L	ND	ND	ND	ND	ND	ND	ND	+	ND
52	-	ND	+	TN	ND	ND	ND	P, LL	D	-	-	+	-	+	-	-	-	-
53	+	+	+	FH, TN	ND	ND	ND	P	Normal	+	-	-	-	+	-	-	-	-
54	-	+	+	TN	ND	+	ND	EF	Normal	-	-	-	-	+	-	+	-	-
55	-	+	+	FH	ND	ND	ND	EF, LL	Normal	-	-	-	+	+	-	-	-	-
JaA	-	+	-	FH, LH, TN, S	ND	ND	ND	P, EF, FE	D	-	+	-	+	+	+	-	-	-
MRS	-	ND	-	ND	-	+	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	-

**Legend:** Patient: Number of code used to identify individuals; Sex: M = male; F = female; Parent of Origin: parent of origin of deletion; MAT = maternal deletion; PAT = paternal deletion; Unknown = unknown deletion; Age in Months: approximate age in months; Scales of Independent Behavior (SIB): SIB BI: Broad Independence score; NA = too low to score; SIB BI AE: Broad Independence score in age equivalent (months); SIB MS AE: Motor skills in months; SIB SC AE: Social/Communication skills in months; SIB PL AE: Personal Living skills in months; SIB CL AE: Community Living skills in months; SIB SS: Support score; Developmental Profile II (DP): DP P: Physical skills in months; DP SH: Self help skills in months; DP S: Social skills in months; DP A: Academic skills in months; DP C: Communication skills in months; DP IQ: Estimated Intelligence score from combined DP scales; Words Spoken: --- = very poor; -- = poor; + = fair; ++ = good; +++ = very good; Hypotonia: + = hypotonia; - = no hypotonia; Ability to walk: --- = very poor; -- = poor; + = fair; ++ = good; +++ = very good; Growth: -- = undergrowth; - = below average; + = average; ++ = above average; +++ = overgrowth; Head Circumference: --- = microcephaly; -- = below average; + = average; ++ = above average; +++ = macrocephaly; Organ Malformations: organ affected by malformation listed; Neurological Scan: + = normal scan; - = abnormal scan; Seizures: + = seizures; - = no seizures; Hearing Defects: + = hearing defect; - = normal hearing; Hands and Feet: S = syndactyly; C = clinodactyly; TN = abnormal tonails; LH = large hands; Recurrent Ear Infections: + = recurrent ear infections; - = no recurrent ear infections; Cesarian Section: + = born by cesarian; - = normal birth; Respiratory Problems at Birth: + = respiratory problems at birth; - = no respiratory problems at birth; Eyes: EF = epicanthal folds; P = ptosis; DS = deep set eyes; FE = full eye lids; LL = long eye lashes; A = abnormal, but not specified; Ears: D = dysmorphic ears; L = long ears; LS = low set ears; S = simple ears; T = preauricular pias and/or tags; P = prominent; A = abnormal, but not specified; High Palate: + = highly arched palate; - = normal arch in palate; Pointed Chin: + = pointed chin; - = no pointed chin; Dental Anomalies: + = dental anomalies; - = no dental anomalies; Full Cheeks: + = full cheeks; - = no full cheeks; Bulbous Nose: + = bulbous nose; - = no bulbous nose; Wide Nasal Bridge: + = wide nasal bridge; - = no wide nasal bridge; Long Philtrum: + = long philtrum; - = no long philtrum; Dolicocephaly: + = dolicocephaly; - = no dolicocephaly; Recurrent Upper Respiratory Problems: + = recurrent upper respiratory problems; - = no recurrent upper respiratory problems





Table 3-6: Summary of Statistical Analysis Comparing Size of Deletion and Clinical Features Observed

Clinical Feature	Coefficient of Correlation	Significance F
SIB BI	0.322	0.166
SIB BI AE	0.609	<b>0.004</b>
SIB MS AE	0.526	<b>0.017</b>
SIB SC AE	0.542	<b>0.014</b>
SIB PL AE	0.545	<b>0.013</b>
SIB CL AE	0.411	0.072
SIB SS	0.182	0.442
DP P	0.426	0.078
DP SH	0.561	<b>0.016</b>
DP S	0.342	0.165
DP A	0.353	0.150
DP C	0.307	0.216
DP IQ	0.090	0.723
Words Spoken	0.162	0.376
Words Recognized	0.056	0.856
Hypotonia	0.302	<b>0.043</b>
Ability to Walk	0.456	<b>0.043</b>
Growth	0.002	0.991
Head Circumference	0.370	<b>0.024</b>
Organ Malformations	0.232	0.168
Neurological Scan Abnormalities	0.144	0.583
Seizures	0.096	0.556
Hearing Defects	0.154	0.357
Chewing Behavior	0.200	0.256
High Pain Tolerance	0.051	0.775
Fleshy Hands	0.259	0.097
Abnormal Toenails	0.169	0.305
Syndactyly	0.032	0.845
Clinodactyly	0.215	0.182
Recurrent Ear Infections	0.236	0.160
Cesarian Section	0.056	0.738
Respiratory Problems at Birth	0.226	0.167
Epicanthal Folds	0.027	0.871
Ptosis	0.246	0.136
Deep Set Eyes	0.169	0.312
Long Eye Lashes	0.007	0.967
Full Eye Lids	0.016	0.925
Preauricular Pits	0.249	0.121
Dysmorphic Ears	0.164	0.299
Abnormal Ears	0.145	0.360
Long Ears	0.199	0.206
High Palate	0.330	<b>0.046</b>
Pointed Chin	0.376	<b>0.026</b>
Dental Anomalies	0.392	<b>0.014</b>
Full Cheeks	0.025	0.888
Bulbous Nose	0.025	0.887
Wide Nasal Bridge	0.275	0.099
Long Philtrum	0.103	0.557

**Legend:** Clinical features are listed in first column as referred to in Table 3-5. The size of the deletion and the clinical features were tested for a correlation using linear regression. Coefficient of Correlation is the multiple R obtained in linear regression analysis. A coefficient of correlation of +1 is a perfect positive correlation, -1 is a perfect negative correlation, and 0 is no correlation. Results were taken to be statistically significant if the F-value was less than 0.05. See appendix C for sample linear regression analysis.



Table 3-7: Summary of Statistical Analysis Comparing Parent of Origin of Deletion and Clinical Features Observed

Clinical Feature	Chi Square Value	P-value	Number of Observations
SIB BI	2.036	0.240	16
SIB BI AE	2.369	0.432	16
SIB MS AE	0.511	0.506	15
SIB SC AE	3.301	0.069	16
SIB PL AE	1.039	0.345	16
SIB CL AE	<b>4.363</b>	<b>0.037</b>	16
SIB SS	0.485	0.560	16
DP P	3.837	0.290	14
DP SH	3.111	0.138	14
DP S	3.111	0.078	14
DP A	2.240	0.320	14
DPC	<b>5.833</b>	<b>0.016</b>	14
DP IQ	2.022	0.551	14
Words Spoken	6.038	0.076	22
Words Recognized	1.667	0.196	10
Hypotonia	0.014	0.905	33
Ability to Walk	2.568	0.386	13
Growth	2.363	0.371	27
Head Circumference	1.762	0.199	26
Organ Malformations	0.005	0.943	27
Neurological Scan Abnormalities	N/A	N/A	14
Seizures	0.068	0.794	30
Hearing Defects	0.008	0.928	29
Chewing Behavior	2.497	0.114	26
High Pain Tolerance	1.501	0.220	26
Fleshy Hands	1.943	0.163	32
Abnormal Toenails	0.003	0.954	30
Syndactyly	1.113	0.291	31
Clinodactyly	<b>3.606</b>	<b>0.058</b>	30
Recurrent Ear Infections	0.622	0.430	27
Cesarian Section	0.714	0.398	30
Respiratory Problems at Birth	0.009	0.925	30
Epicanthal Folds	1.050	0.306	28
Ptoisis	<b>3.996</b>	<b>0.046</b>	28
Deep Set Eyes	0.049	0.825	28
Long Eye Lashes	0.078	0.780	26
Full Eye Lids	0.207	0.649	28
Preauricular Pits	1.530	0.216	32
Dysmorphic Ears	1.891	0.169	32
Abnormal Ears	1.971	0.160	32
Long Ears	1.176	0.278	31
High Palate	0.124	0.724	28
Pointed Chin	1.187	0.276	27
Dental Anomalies	0.373	0.542	30
Full Cheeks	1.418	0.234	26
Bulbous Nose	1.222	0.269	26
Wide Nasal Bridge	0.123	0.726	27
Long Philtrum	0.123	0.726	27
Dolicocephaly	0.000	1.000	28
Recurrent Respiratory Infections	1.209	0.271	29

**Legend:** Clinical features are listed in first column as referred to in Table 3-5. Chi Square values are obtained using contingency tables for the expected values. P-values show the statistical support for the null hypothesis that the clinical features and the parent of origin of the deletion are independent variables. With a p-value of 0.05 or less the null hypothesis was considered not to be supported by the data, and indicating that the two variables may be dependent. See appendix D for sample analysis.



## Chapter 4: Discussion

### Characterization of the human ortholog of rat *Shank3*:

Through BLAST searches with a cDNA partially characterized by A. Wong, the gene *SH3 and multiple ankyrin repeat domains 3 (SHANK3)* was characterized. This gene had been newly identified in rat (also called *ProSAP2*, and *Spank-2*) (Boeckers, et al, 1999a, and Boeckers, et al, 1999b). There have been two other closely related genes identified in rat, *Shank1* (also called *synamon*, *SSTRIP* and *Spank1*) (Naisbitt, et al, 1999), and *Shank2* (also called *ProSAP1*, *Spank3* and *CortBP1*) (Boeckers, et al, 1999a and Boeckers, et al, 1999b). Through studies in rat these proteins have been localized to the postsynaptic density of excitatory synapses, and are an important in linking the surface receptors and ion channels to the cytoskeleton (figure 4-1).

The proteins of the Shank family have at least six different protein interacting domains. The PDZ domain is found in many other proteins most of which are localized to the synaptic region of neurons (reviewed in Craven and Bredt, 1998). The PDZ domain of the Shanks binds to GKAP/SAPAP. The GKAP/SAPAP proteins bind to PSD-95 which in turn interacts with the NMDA receptors located on the surface of the postsynaptic membrane (Boeckers, et al, 1999b; Naisbitt, et al, 1999). NMDA receptors are responsible for receiving the glutamate signals released from the presynaptic membrane. The HOMER-binding domain of the Shank family binds to a homodimer of HOMER which connects to the metabotropic glutamate receptors on the postsynaptic membrane (Tu, et al, 1999). The SH3 binding domain of the Shank proteins joins the NMDA and the metabotropic glutamate receptors to the actin cytoskeleton through



Cortactin (Boeckers, et al, 1999a, and Naisbitt, et al, 1999). The Sterile Alpha Motif (SAM domain) is responsible for the homodimerization of Shank family members in a tail to tail manner (reviewed in Ehlers, 1999). The ankyrin repeat domain is responsible for the interaction of the Shank family of proteins with Sharpin. The interaction of the Shanks with Sharpin results in a network of proteins in the postsynaptic density formed by linking the Shanks together in a head to head fashion (Lim, et al, 2001). The ankyrin repeat domain has not been found in any characterized isoforms of Shank2. The SH3 domain is found in all three Shank proteins, however, its interaction with other postsynaptic density proteins has not yet been characterized. Thus the Shank proteins are a key structural component in forming a continuous network of proteins in the postsynaptic density that join the two main types of receptors on the postsynaptic membranes to the cytoskeleton and the signal transduction pathways.

Since the Shank structural proteins interact with multiple other proteins to form the continuous network in the postsynaptic density, any alteration of the relative proportions of such a proteins could have devastating effects on the synapse formation and function. Thus the deletion of *SHANK3*, as in the patients with the 22q13 deletion syndrome, would likely contribute to the phenotype.

### **SHANK3 Transcript Analysis:**

The transcripts of *SHANK3* in humans have been characterized in this thesis (figure 3-2). *SHANK3* is predicted to be composed of 27 exons that spans a genomic region of approximately 60kb. Two of the regions were not characterized in the human cDNAs by A. Wong (1998). The sequence of the first 162 bases of the rat cDNA is





highly similar to a region in the human genomic sequence immediately upstream of the most 5' cDNA characterized by A. Wong. Since this region is part of the rat cDNA and is present in the human genomic sequence there is a strong possibility that it is part of the first exon of the human transcript.

A second region of the rat cDNA, that corresponds to exon 11 in humans, has not been obtained in a human cDNA and is not present in the genomic sequence that surrounds the region where this exon should be (figure 3-2). This region remained as a sequencing gap until May 2000 when this gap suddenly disappeared from the database and the two sequences on either side were joined together. Three possibilities exist: 1) The region containing this exon was deleted out of the cosmid that was sent for sequencing. Several small deletions have been identified in other genomic clones that are part of the sequencing project (McDermid, personal communication). However, initially a gap was localized to this spot, and this possibility does not explain the sudden disappearance of the gap. The presence of a deletion in the cosmid could be determined by comparing restriction digests of the cosmid to restriction digests of the genomic DNA by Southern blot analysis. 2) This region is extremely G/C rich and/or repetitive and is therefore extremely difficult to sequence. The sequence of the region in the rat cDNA is 97% G/C, and the human sequence is expected to be similar. High G/C content in a region of DNA makes sequencing difficult. This is a good possibility since several other regions of the same cosmid were difficult to sequence. It might be worthwhile to try other sequencing strategies to verify the genomic sequence of this region such as using manual sequencing rather than automated sequencing and altering the different



components of the sequencing reaction including the amount of enzyme and the concentration of DMSO. 3) This exon is not present in humans. The exon is 60bp in size, and therefore would not affect the open reading frame. No protein interacting domains have been identified within this exon, so the function of the protein may not be affected by the lack of this exon. Several attempts were made to try to obtain this exon. RT-PCRs from exons nearby were tried on heart, muscle, and kidney templates. Products were obtained from these RT-PCR reactions, but none of them contained exon 11 or 12. Exon 12 was shown to be part of some transcripts through other RT-PCR reactions that started in exon 12. Thus, alternative splicing of this gene does exist. Exon 11 may not have been obtained in an RT-PCR product because the wrong tissue or the wrong time point was used in the RT-PCR experiments. In rat the *Shank3* gene shows high levels of expression in brain that is undergoing active synaptogenesis (Boeckers, et al, 1999b). Thus RT-PCR to try and obtain this exon should be tried on brain tissue from brain that is undergoing synaptogenesis, such as neonatal or adult brain. Specific regions of neonatal or adult brain, such as cerebellum, should be used in RT-PCR since from the Northern blot data, different regions of the brain show different expression patterns of *SHANK3*. Another possibility for the absence of an RT-PCR product containing this exon is that there is preferential amplification of other transcripts in the tissues used. From the Northern blot analysis expression of the 7.5kb brain predominant transcript is also expressed in kidney, but at much lower levels. The amplification with primers in exons on either side of exon 11 may preferentially amplify the region from one of the other transcripts which are more abundant in this tissue. A final possibility is that this exon is



very G/C rich and is impossible to synthesize cDNAs that contain this exon, or to amplify this exon by PCR. Sequencing was tried unsuccessfully in both human cosmid and genomic DNA using primers made to the sequence of the rat cDNA. Sequencing reactions using these primers may not have worked because there may have been enough divergence between the rat and human sequence such that the primers did not anneal to the human sequence. The sequencing may not have worked since the sequencing enzyme may not be able to process the high percentage of G/C observed in this region. Primers were also designed flanking the region of the suspected exon, and PCR was carried out using genomic DNA as a template. A product was obtained but all attempts to sequence this product failed. This may be due to a high G/C content of this region. It is therefore inconclusive at this stage as to whether exon 11 is part of the human cDNA sequence or not.

Exon 22 was identified by Wong (1998) in a cDNA isolated from a library (figure 3-2). This exon has not yet been obtained in any RT-PCR products. This exon is also not present in the sequence of the rat cDNA. Thus exon 22 may not be a true exon of the *SHANK3* gene, or may be part of one of the alternatively spliced transcripts of this gene. Further experiments using RT-PCR should be tried in different tissues to determine if this exon is part of one of the transcripts of *SHANK3*.

In the analysis by Wong (1998) three alternative 3' ends of *SHANK3* were identified through exon prediction and cDNAs. This indicates that there are several different transcripts of *SHANK3*, which was confirmed by observing four different bands on Northern blots. RT-PCR between exon 23 and two of these 3' ends verified the



presence of these 3' ends in transcripts of this gene (figure 3-2). The third 3' end, exon 26, was never obtained in an RT-PCR product, or through RACE reactions. This exon contains the SAM domain of the SHANK3 protein and is thus a key domain for the function of this protein in the synapses. Exon 26 may be specific to transcripts that are found in neurons, and since infant/adult brain tissue was not available for analysis, perhaps this exon was not detected in heart, muscle and kidney. Further experiments should be carried out using various regions of the infant/adult brain as an RNA template for the RT reaction.

The two largest transcripts, 8 and 10kb, observed in the Northern Blot likely represent the alternative 3' ends, exons 24 and 27 respectively (figure 4-2). Exon 11 is not likely contained in these two transcripts of *SHANK3* since RT-PCR on tissues that show expression of these two transcripts did not give any products containing exon 11. It is not clear if exon 22 is part of these transcripts or not, no RT-PCR products were obtained that contain this exon. However, when the sizes of the possible transcripts are added up and compared to the size observed on the Northern blots, it is possible that exon 22 is part of these two larger transcripts. Both of these large transcripts likely include the 2kb exon 23 but likely do not have the SAM domain, meaning that the resulting proteins would not form homodimers. The content of these two transcripts are likely similar in the exon content with the exception of the 3' end. The 2kb truncated transcript likely does not contain all of the key protein interacting domains of SHANK3. Only the first half of *SHANK3* is predicted to be contained in this transcript (exons 1 to 13, without exons 11 and 12). The function and the 3' end of this transcript are unknown. The 7.5kb brain





predominant transcript is predicted to be the same as the rat cDNA obtained for *SHANK3*. This transcript would likely contain exons 1 through 21, and exons 23 and 26. These exons correspond to the rat cDNA and produce a band that is similar in size to the band observed on the rat Northern blot analysis (Boeckers, et al, 1999b). Identifying the different isoforms of *SHANK3* will be a complex and difficult task that would require RT-PCR analysis to try and join other regions, probing Northern blots with other regions of the gene and trying to obtain cDNAs that contain more of the gene.

The other members of the Shank proteins have been shown to have alternative splicing. Boeckers, et al (1999a) found at least three different processing variants of the gene Shank2 which affected the N-terminal end. The alternative splicing and alternative 3' ends would result in multiple isoforms of the protein with different combinations of the protein interacting domains and thus different functions. Different isoforms of the gene may be expressed in different tissues, perhaps coding for very specific functions in these tissues.

### **Expression of SHANK3:**

In humans *SHANK3* has a complex expression pattern. Four different transcripts are detected at different levels in the tissues tested by Northern blot analysis (figure 3-5). The 10kb transcript is present in all of the tissues tested at lower levels, but shows a much higher level of expression in the cerebellum. The lower level of expression of the 10kb transcript in the other tissues may be due to the inefficient transfer of the larger RNA molecules. However, since strong expression is observed in the cerebellum, the inefficiency of transfer does not seem to adequately explain this difference. The 8kb



transcript is present in all of the tissues tested at low levels, but shows a higher level of expression in the brain and heart. Interestingly the 8kb transcript is not expressed in the cerebellum. Therefore the 8kb transcript may have a specific function in these tissues, that differs from the function characterized in the rat. The most interesting transcript is 7.5kb and corresponds in size to the only transcript observed on the rat Northern blots (Boeckers, et al, 1999b). It is highly expressed in most parts of the brain, but although it is brain predominant, it is also at lower levels in the kidney. The expression of *SHANK3* observed in other non-neurological tissues may indicate that the SHANK3 protein can also function in the formation of a variety of other sites of cell-cell interaction. A good candidate gene for the neurological features associated with the 22q13 deletion syndrome would be expected to show strong expression in the brain and nervous system tissues. Since *SHANK3* shows strong expression of the 7.5kb transcript in most of the tissues of the nervous system, it is a good candidate gene for the neurological features of the 22q13 deletion syndrome.

The expression of *Shank3* has been examined in several model organisms by Northern blot analysis. In rats, a single transcript was observed, 7.5kb that was only found in the brain (Boeckers, et al, 1999b). In pigs a single transcript was found, approximately 7.5kb, expressed predominantly in the brain, but at lower levels in the lung (figure 3-7). Several different parts of the brain showed expression of this transcript. The mouse Northern blot showed two transcripts 7.5kb and 8kb that are predominantly expressed in the brain, with a very low level observed in the lung. It is interesting that both the pig and the mouse show expression of *Shank3* at low levels in the lung, whereas



lung expression is not mentioned in rat. Boeckers, et al (1999b) do not mention which tissues they used in their rat Northern blot, perhaps lung was not included. The presence of the 8kb transcript observed in mouse is unusual, neither of the other two model organisms show this transcript. Perhaps a probe that would only recognize the 7.5kb transcript (3' end) was used in the Northern blot analysis carried out in the rat (Boeckers, et al, 1999b). However, this does not explain the absence of this transcript on the pig Northern blot. The brain specific expression pattern in rodents compared to the broad expression in humans is not a unique situation. Necdin, a candidate gene in the Prader-Willi region on chromosome 15, shows a similar pattern of expression: brain specific in rodents and a broad tissue range in humans (MacDonald, and Wevrick, 1997). This poses an interesting question, as to why the expression pattern in humans is different from the other mammals studied. It is possible that *SHANK3* is involved in higher brain function, such as speech in humans and thus is not present in these lower mammals. However this does not explain the presence of the other transcripts of *SHANK3* in many different tissues in humans.

Since *SHANK3* is a member of a family of proteins that are similar in structure and sequence, the redundancy of function of these proteins may enable compensation for the loss of one of these family members. However mRNA *in situ* hybridization of *Shank2* and *Shank3* have shown slightly different patterns of expression in rat brain (Boeckers, et al, 1999b). *Shank2* has been localized to several parts of the brain including high levels of expression in the cerebellum, hippocampus, cerebral cortex, central gray matter and the olfactory bulb. Immunohistochemical localization of the protein showed





expression in the cerebral cortex, the molecular layer of the cerebellum, basal ganglia, hippocampal formation and thalamic nuclei (Boeckers, et al, 1999a). Specific punctate patterns were noticed by Redecker, et al (2001) within the lining of the cerebral ventricles, the aqueductus cerebri and the central canal of the spinal cord. The cells that were positive for Shank2 were identified as tanycytes by vimentin antibodies. Shank2 expression was also observed in several different types of endocrine cells including pancreatic islet cells. Immunohistochemistry also showed Shank2 expression in the hepatic bile ducts epithelium, tubule system of the kidneys and the bronchiolar tree of the lungs, which may explain the weak expression observed in human kidney. Using mRNA *in situ* hybridizations Shank3 appears to co-localize with Shank2 in the cerebral cortex and the hippocampus (Boeckers, et al, 1999b). However in the cerebellum Shank3 is expressed only in the granular cell layer, whereas Shank2 is expressed primarily in the Purkinje cells. Thus this family of proteins may have a similar function but the complementary expression observed between Shank2 and Shank3, particularly in the cerebellum, indicates that their function may not be redundant.

Similar immunohistochemical localization should be carried out on samples of human brains. Antibodies specific to each of the different Shank proteins should be designed and the differences in the cellular expression pattern should be examined. The immunohistochemical localizations should be carried out on human brain samples that are undergoing active synaptogenesis. This means that the ideal brain for this analysis is a neonatal brain (preferably within the first year of life); unfortunately this specific type of sample is difficult to obtain. Different regions of the brain should be analyzed to





determine if the expression pattern of the Shanks is specific to different regions. Special attention should be focused on the cerebellum, since a difference of Shank expression is observed in this brain region in the rat, and the various regions of the brain that are involved in speech production. Immunohistochemical experiments on different regions of human brain samples would allow for confirmation that the different Shank proteins show different expression patterns at the cellular level.

A recently published abstract (Jee, et al, 2001), identified the Shank homologue in *C.elegans*, which was also partially characterized by Wong (1998). This homologue, *shn-1*(C33B4.3), shows approximately 40% similarity to the Shank proteins over 1000 amino acids. The similarity is higher in the PDZ domain and the ankyrin repeat domain. The promoter region of this gene was fused to a GFP construct to observe expression of the Shank homologue. GFP was observed primarily in the pharyngeal muscle, head sensory neurons, nerve cord and the tail region. An antibody made to the ankyrin repeats found in rat Shank1 was used for whole mount immunohistochemistry and confirmed the GFP expression analysis. The authors commented that the ability to detect *shn-1* with an antibody to the rat homologue highlights the conservation of the Shank proteins between species, even species that are quite evolutionarily removed. Studies are underway by these authors to determine the effect of this gene in the *C. elegans* development. RNA interference experiments as well as deletion mutation screens are being used to determine the effect of the loss of this gene.

The effect of the Shank3 gene on development in mammals is underway. A knock-out rodent is being created (Boeckers, et al, personal communication). The effect



of knocking out this gene in rodents would be an interesting experiment. If hemizygosity of *SHANK3* causes developmental delay and delay of expressive speech in the 22q13 deletion syndrome, difficult behavioral assays on the rodents would be necessary. Detecting developmental delay in rodents is difficult, and the study of speech is impossible. Thus, it may be difficult to model the phenotype observed in the 22q13 deletion syndrome by knocking out the *Shank3* gene in rodents.

Bonaglia, et al (2001) also identified *SHANK3* based on the similarity to the rat homologue. In this study, none of the alternative 3' ends were identified since they were not part of the rat cDNA. Exons 11 and 22 are also not identified by this study. The Northern blot analysis by this group shows slightly different results than those obtained in this thesis. The smallest transcript is 3kb (rather than 2kb in this thesis) and shows expression at different levels in several different tissues. This study also determined that the two middle sized transcripts of 7kb and 8kb (corresponding to 7.5kb and 8kb in this thesis) are brain specific, with lower levels in the medulla and spinal cord. In this thesis, only the 7.5kb transcript was shown to be brain predominant; the 8kb transcript showed expression in most tissues tested. Bonaglia, et al (2001) also showed that the largest transcript is only expressed in cerebellum tissue, whereas this thesis showed that the highest level of this transcript is observed in the cerebellum, but that it is present at lower levels in other tissues.

#### Future Work:

The characterization of the *Shank* proteins in rodent models is well underway. The effect of knocking out one of the *Shank* genes should provide information regarding



the role of this gene in brain development and functioning. However, the rodent models will not model the role of SHANK3 in higher functions such as speech. Thus, human studies of the *SHANK3* gene are necessary. Since *SHANK3* in humans has a different expression pattern than rodents at the mRNA level, further investigations on the expression pattern and possible specialized functions of SHANK3 need to be conducted. Northern blots need to be probed with the different family members of the Shanks to determine if there is a difference in the mRNA expression. It would also be useful to probe Northern blots with different regions of the Shank genes to determine if there is a specific expression pattern for the different regions and to work out which exons make up the different transcripts. Immunohistochemical localization of SHANK3 would be useful to determine what types of cells SHANK3 is produced in. Other tissues which show expression of *SHANK3* at the mRNA level may not show expression of SHANK3 at the protein level. The SHANK3 protein may only be found in the neurons of these tissues, and not the other cell types. This possibility does not seem likely since expression is detectable on Northern blot analysis, and neurons are not found in large quantities in these other tissues. The role of the several different transcripts observed in humans needs to be characterized, especially since many of these transcripts are not specific to neurological tissues. The protein expression pattern of SHANK3 may give a better understanding of how this gene might be responsible for the neurological abnormalities observed in the 22q13 deletion syndrome.

### **Analysis of Patient Deletions:**

In order for SHANK3 to be considered as a candidate gene for the neurological



abnormalities associated with the 22q13 deletion syndrome, it must be hemizygous in all individuals with the syndrome. In order to test this, all patients were analyzed with the minisatellite D22S163. All patients successfully analyzed showed a single allele at this locus (51/52 patients) with the exception of one patient that will be discussed below. The amount of DNA in some of the patient samples was not enough to produce a visible band on the Southern Blot, and therefore these patients appear as ND in table 3-2 (6/57 patients, numbers: 2, 28, 38, 40, 44 and JaA). In 44% (12/35) of the cases where parental DNA was available, the minisatellite was uninformative. This indicates that D22S163 cannot be used efficiently to confirm terminal deletions. If parental DNA was not available or the minisatellite was uninformative, then FISH analysis was used to confirm the deletion of *SHANK3*. Probes that include *SHANK3*, as well as more telomeric probes were used to confirm the deletion of *SHANK3* in 16 cases. Metaphase cell preparations were not available on 13 patients in this study and therefore the deletion of *SHANK3* could not be confirmed. Data on 10 patients with metaphase cell preparations was not available for the deletion of *SHANK3*. Several of these patients had poor metaphase spreads and thus repeating the experiment would not likely give different results. With some of the metaphase spreads the probe did not hybridize and repeating the FISH on these patients would be useful. Deletion of *SHANK3* was confirmed in 33 out of the 56 patients with the 22q13 deletion syndrome. Thus, the majority of the deletions of 22q13 have a confirmed deletion of *SHANK3* providing additional support for its involvement in the neurological features observed in the 22q13 deletion syndrome. None of the patients in this study, with the exception of MRS, showed two copies of *SHANK3*.





Microsatellites provided an efficient way to approximate the size of the deletions in a large number of patients especially if parental DNA samples were available. By using microsatellites throughout the region, the use of the more difficult technique of FISH could be minimized. The upper limit of all of the deletions was determined by the presence of two alleles at a particular microsatellite locus or heterozygosity of the locus. The informativeness of the microsatellites showed considerable variation. Microsatellites such as D22S1159 (59% uninformative) proved to be not as useful in the analysis of patient deletions and therefore should not be used unless absolutely necessary. The region containing the breakpoint of the deletion was completely identified by microsatellite analysis in 21 out of 57 cases. However, microsatellite analysis was not enough to determine the size of the deletion in most cases and FISH had to be used.

FISH allowed for confirmation and/or better refinement of the deletion in patients without parental DNA or patients with large regions of uninformative microsatellites. Cosmids were available throughout the region. FISH was used to narrow down the region containing the deletion breakpoint in 22 out of the 57 patients. One problem with FISH is the requirement for metaphase cell preparations from either blood or a cell line. Unfortunately metaphase cell preparations were not available for 16 of the patients with the 22q13 deletion syndrome.

By combining the information obtained from D22S163, and the microsatellites the parent of origin of the deletions were obtained in 41/57 cases. There were significantly more paternal germline deletions observed than maternal germline deletions (Chi square value 4.9 with one degree of freedom and a p value of 0.025). However,



since both types of deletions are abundant, there is little if any significant imprinting effect in this region. Several studies indicate no evidence for imprinting on chromosome 22. For example three cases of maternal uniparental disomy resulting from a Robertsonian translocation (t22;22) showed no abnormal phenotype (Schinzel, 1994). Thus imprinting likely does not have a large contribution to the phenotype of the 22q13 deletion syndrome.

The deletions showed considerable variation in size from 130kb to over 10Mb in size. At the beginning of this project patient 1/NT was by far the smallest characterized terminal deletion of 22q, and there were some doubts that his phenotype was caused by the deletion. 1/NT's clinical features were not as severe as those observed in the other patients with the 22q13 deletion syndrome (Flint, et al, 1995). 1/NT has only mild developmental delay and absence of expressive speech with no dysmorphic features commonly observed with the 22q13 deletion syndrome. The next smallest deletion characterized before this study was case 1 published by Doheny, et al (1997) (figure 4-3). This patient has developmental delay, delay of expressive language, hypotonia and several dysmorphic features associated with the 22q13 deletion syndrome. This patient's deletion is approximately 2Mb in size and is deleted for the probe D22S39 (Oncor control probe for DiGeorge/VCF syndrome). Due to the large difference in size between the two smallest characterized deletions and the mildness of 1/NT's phenotype, it was speculated that 1/NT's phenotype might be due to some other undetected genetic cause. In this thesis five patients (2, 3, 4, 5, and 6) were characterized that have a smaller deletion than case 1 (Doheny, et al, 1997). Patient 2 is the result of a translocation and will be



discussed below. The deletion in this patient is a little larger than the NT deletion. Patient 3 is not deleted for the probe D22S39, but more distal probes are deleted in this patient. The upper limit of the breakpoint containing region in patient 3 is currently defined by the cosmid cN113F5. The location of this breakpoint is within the gene rich region of the terminal region. Patient 4 is not deleted for the probe D22S39, whereas patients 5 and 6 are both deleted for this probe and all three are similar in size to case 1 published by Doheny, et al (1997). None of these four patients provide any insight into other genes that may be involved in the 22q13 deletion syndrome since their breakpoint is within the gene poor region and no genes can be eliminated. Patient 2 shows the typical clinical features of the 22q13 deletion syndrome including developmental delay, delay of expressive speech and hypotonia. Thus patient 3 is of particular importance in defining the important region for this deletion syndrome, since the breakpoint of this patient falls within the gene rich region adjacent to the telomere.

### **Analysis of Patient Phenotype:**

The clinical features of the patients characterized in this study are similar to previously published studies. All of the patients were found to have the main neurological findings of developmental delay and delay or absence of expressive speech. Unfortunately detailed clinical information was not available on all of the patients involved in this study. However the large amount of clinical data and large patient set allowed statistical analysis tools to be used to detect correlations between the size of the deletion and the clinical features present, as well as between the parent of origin of the deletion and the clinical features present.



In comparing the size of the deletion and the various clinical features, several correlations are apparent. The age equivalent Scales of Independent Behavior Broad Independence test shows a strong, significant correlation (correlation coefficient of 0.609). Several other IQ features also show a significant correlation, including: Scales of Independent Behavior Motor Skills (0.526), Social/Communication Skills (0.542), Personal Living Skills (0.545), and Developmental Profile II Self Help (0.561). Several of the other clinical features also show a significant positive correlation, including: hypotonia (0.302), ability to walk (0.456), head circumference(0.370), high palate (0.330), pointed chin (0.376), dental anomalies (0.392), and recurrent respiratory infections (0.323). Therefore the larger the deletion, the higher the incidence or the more severe some of the clinical features are. No negative correlations were observed in the statistical analysis of the size of the deletion and the clinical features. The IQ features that show a significant positive correlation indicate the possibility of additional genes with involvement in developmental delay being located in the proximal region of 22q13. However this is not too surprising since many genes are involved in brain development and therefore an additive effect may be observed in the larger deletions. Genes in this region may also contribute to hypotonia, and a change in muscle tone would affect the ability to walk. The clinical features of the high palate, dental anomalies, and pointed chin are all a function of abnormalities of the same facial region, the jaw. The more prominent occurrence of these features in the larger deletions suggests the possibility that a gene in proximal 22q13 may be responsible for jaw formation. Also, a gene in proximal 22q13 may contribute to the increased susceptibility to upper respiratory





infections of these patients. Several different types of genes may increase the susceptibility including immune function genes, and other genes involved in lung functions. No critical region can be defined for each of these clinical features since both large and small deletions show these clinical features. However there may be contributing factors to the different clinical features since many features have a higher incidence in the larger deletions. Since these are likely contributing factors and not solely responsible for the features, the development of these features must be multifactorial with the involvement of many other genes.

The statistical analysis also showed the lack of any significant correlation to the size of the deletion for several of the clinical features. Interestingly some of the different components of the IQ did not show any correlation, these include the Scales of Independent Behavior Broad Independence (not age equivalent)(0.322), Support Score (0.182), the Developmental Profile II Physical Skills (0.426), Social Skills (0.342), Academic Skills (0.353), Communication Skills (0.307), and Estimated Intelligence Score (0.090). The fact that these different parts of the IQ scores do not show correlation of the severity to the size of the deletion indicates that the genes controlling these parts of IQ must be deleted in all of the deletions. This also is seen by the fact that neither the words spoken (0.162) nor words recognized (0.056) correlate to the size of the deletion. Therefore since *SHANK3* is deleted in all 22q13 deletions tested, it may be responsible for these components of the IQ score. Interestingly growth did not show a correlation to the size of the deletion, but head circumference did. This suggests that a gene in the proximal region is responsible for regulating head size, but that this gene is in no way



connected to the growth of the patient. Organ malformations, seizures and hearing defects were clinical features that were expected to show a correlation to the size of the deletion. It would make sense that the larger the deletion, the more problems there are with development. However no significant correlation was detected for any of these three clinical features. Many of the mild dysmorphic features associated with the 22q13 deletion syndrome did not show any correlation to the size of the deletion.

The presence or absence or the severity of the clinical features of the 22q13 deletion syndrome, were also analyzed for a correlation to the parent of origin of the deletion (40/56 patients). For many of the clinical features, no correlation to the parent of origin is observed. However for some of the features statistical tests show possible correlations. Clinodactyly occurs more frequently in maternal germline deletions than in paternal germline deletions (p-value 0.058). The opposite is true for ptosis (p-value 0.046), paternal germline deletions have a higher incidence than maternal germline deletions. Although the IQ scores show little correlation between the parent of origin and the severity, a possible effect is observed in the Scales of Independent Behavior Social/Communication Skills (p-value 0.069), Community Living Skills AE (p-value 0.037) and Developmental Profile II Communication Skills (p-value 0.016) portions, where paternal germline deletions appear more severe than the maternal germline deletions. A correlation to the parent of origin would be suggestive of a mild imprinting effect on genes in this region. All of these correlations are made on relatively small sample sizes, the largest being 33 patients. Larger sample sizes may show that these correlations were just an anomaly of the sample in this study. If these correlations are



valid, then there may be some genes in this region that show differential expression of alleles depending on the parent from which the allele was inherited. There may be a gene preferentially expressed from the maternal allele that influences digit formation or genes preferentially expressed from the paternal allele that may contribute to facial or brain development. Since the clinical features that show correlations are observed in both maternal and paternal germline deletions, one cannot say that there are fully imprinted genes in this region, but that there may be differential expression.

### **Analysis of Genes in 22q13 Deletion Region:**

Since there is some effect of the size of the deletion on the clinical features observed, the more proximal region of 22q13 needs to be examined for genes that might contribute to these phenotypes. By combining the information available on the Sanger Center web site ([www.sanger.ac.uk](http://www.sanger.ac.uk)), NCBI's UniGene site ([www.ncbi.nlm.nih.gov/UniGene](http://www.ncbi.nlm.nih.gov/UniGene)) and Online Mendelian Inheritance in Man (OMIM; [www.ncbi.nlm.nih.gov/omim/](http://www.ncbi.nlm.nih.gov/omim/)), one can search for potential candidate genes *in silico*. For instance the expression pattern of all genes and predicted genes can be approximated based on the tissue sources of ESTs and "Serial Analysis of Gene Expression" (SAGE) Tag to Gene Mapping. OMIM provides information on mutations and genetic disorders involving known genes. The region of the largest 22q13 deletion syndrome patients has therefore been analyzed to determine if there are any potential candidate genes for other clinical features observed in the syndrome.

The distribution of the genes in 22q13 is unusual (figure 4-3). There is a large cluster of about 30 genes (including *SHANK3*) located at the most distal end of 22q13,





just proximal to the subtelomeric repeats. Centromeric to this gene dense region there is a large gap, of approximately 2.5Mb, where there are only a few predicted genes. The region proximal to this gene poor region has an even distribution of genes. Many of these genes have been well characterized and have already been implicated in genetic disorders.

Most of the deletions in patients with the 22q13 deletion syndrome (except 1/NT, patient 2 and 3) result in the deletion of the gene rich region adjacent to the subtelomeric repeats. Many genes in this region have been well characterized and are not likely candidate genes. As discussed in the introduction *RABL2B* and *Acrosin* are not good candidate genes for the neurological abnormalities. Several of the other genes encode enzymes, or other proteins that are not sensitive to dosage. For instance, *ARSA* encodes an enzyme involved in hydrolysis of cerebroside sulfate (Kreysing, et al, 1990), and is responsible for autosomal recessive metachromatic leukodystrophy. Patients with metachromatic leukodystrophy show a progressive demyelination that leads to neurological symptoms and death (OMIM: 250100). *ECGF1* catalyzes the phosphorolysis of thymidine to thymine (Nishino, et al, 2001) and is responsible for autosomal recessive mitochondrial neurogastrointestinal encephalomyopathy. These patients have a variety of gastrointestinal problems including abdominal pain and diarrhea, muscle atrophy and mild neurological features. *SCO2*, a member of the cytochrome c oxidase, is responsible for autosomal recessive fatal infantile cardioencephalomyopathy (Papadopoulou, et al, 1999). Patients with mutations in *SCO2* show hypertrophic cardiomyopathy, lactic acidosis and gliosis, and reductions in COX activity. Since all three of these diseases show no phenotype in the heterozygote, these





genes would not be candidate genes for the clinical features of the 22q13 deletion syndrome.

Two other genes in the critical region have not been well characterized and have not yet been associated with a genetic disorder. The *PLXNB2* gene was isolated from a screen for genes differentially expressed in malignant vs benign brain tumors (Shinoura, et al, 1995) suggesting a possible role in tumor malignancy. The *MOV10L1* gene was identified through a screen for genes expressed in the spermatogonia (Wang, et al, 2001). *MOV10L1* is predicted to function as an RNA helicase. Neither of these genes appear to have a major impact on the clinical features observed in the 22q13 deletion syndrome.

Several of the genes in this region may have an impact on the hypotonia which can result from abnormalities of the nervous system or muscles. *CPT1B* is a protein that is primarily expressed in muscles and is involved in the transport of long chain fatty acids into the mitochondria (Yamazaki, et al, 1996). The lack of this protein may have an impact on the proper function of the muscle cells. *SBF1* has a protective function preventing substrate dephosphorylation (OMIM: 603560). Experiments by Cui, et al (1998) showed that this gene is important molecular component of growth control. Ectopic expression has an effect on the differentiation of myoblast cells in vitro, and thus the lack of expression of this gene may not allow for proper differentiation of muscle cells. The two MAPK genes, *MAPK11* and *MAPK12*, are mitogen-activated protein kinases and are involved in the transduction of extracellular signals (OMIM 602399). Northern blot analysis of the *MAPK12* gene shows high levels of expression in the skeletal muscle (Li, et al, 1996). *MAPK12* functions as a signal transducer in the



differentiation of myoblasts into myotubes (Lechner, et al, 1996). It is possible that any of these genes affecting muscle function may contribute to the hypotonia associated with the 22q13 deletion syndrome.

The *MLC1/WKLI* gene was originally identified from a myeloid cell line (Nomura, et al, 1994). Northern blot analysis of this gene shows high expression in the brain with lower levels in a variety of other tissues. Leegwater, et al (2001) found mutations in this gene in individuals with megalencephalic leukoencephalopathy (MLC) with subcortical cysts. This disorder affects the white matter of the brain and leads to slow progressive ataxia and spasticity (Van der Knaap, et al, 1995). Epilepsy is also observed in many of the patients diagnosed with MLC. This disorder is believed to be inherited in an autosomal recessive pattern, so carriers of one mutant copy are normal. Since this is inherited in an autosomal recessive manner, the patients with the 22q13 deletion syndrome are not expected to show MLC.

Meyer, et al (2001) also analyzed the *MLC1/WKLI* gene as a candidate gene for catatonic schizophrenia, which some linkage studies map to 22q13. This form of schizophrenia is characterized by psychosis, psychomotor disturbances, increased anxiety, impulsivity, aggressiveness, and hallucinations and delusions. Through sequencing *MLC1/WKLI* in affected individuals, a C to A transversion was identified that results in a missense mutation. This transversion creates a restriction site for SfaNI allowing for easy screening of individuals for this mutation. This mutation was detected in all affected individuals as well as all of the obligate carriers. This result was interpreted to show incomplete penetrance, or a mild, undiagnosed, phenotype that is inherited in an



autosomal dominant manner. *MLC1/WKL1* is hypothesized to be a non-selective neuronal cation channel that is related to the shaker potassium channels. This gene is most highly expressed in the nigro-striatal motor and the mesolimbic systems in the brain. Since this gene has been characterized to be involved in an autosomal dominant form of catatonic schizophrenia, it may have an impact on the phenotype of the 22q13 deletion syndrome. All but 2 of the patients studied are hemizygous for the *MLC1/WKL1* gene. Thus, it is feasible that some of these patients may show some of the clinical features associated with this catatonic schizophrenia. However, due to the severe developmental delays and the delay or absence of expressive speech in these individuals, it would be hard to ascertain if any of the clinical features of catatonic schizophrenia are present. A recent abstract screened 200 cases of schizophrenia for the same mutation identified in Meyer, et al (2001), but failed to detect this missense mutation (McQuillin, et al, 2001). This study did identify an insertion/deletion of the *MLC1/WKL1* gene, but this insertion/deletion was also found in controls at Hardy-Weinberg equilibrium. Therefore the evidence that *MLC1/WKL1* gene is responsible for schizophrenia is somewhat controversial.

Several other genes that are found outside of the critical region for the 22q13 deletion syndrome may have significant contributions to the clinical features observed in the larger deletions. The *CELSR1* gene is a membrane receptor that is involved in contact mediated communication, and has several motifs that mediate protein interactions (Hadjantonakis, et al, 1997; Hadjantonakis, et al, 1998). *CELSR1* is primarily expressed in nervous system tissues, including the brain, neural tube, spinal cord and eye and may



therefore contribute to the neurological abnormalities when deleted. Underexpression of *CELSRI* might contribute to the lower IQ score observed in patients with larger deletions. *PKDREJ* shows expression in the testis by Northern blot analysis, however it is closely related to a gene involved in polycystic kidney disease (Hughes, et al, 1999). Several patients with the 22q13 deletion syndrome have kidney abnormalities, and a few have cystic kidney disease. *PKDREJ* is deleted in all of the patients with some type of kidney abnormality. It seems possible that hemizyosity for *PKDREJ* may predispose these patients to polycystic kidney disease although most are normal to date.

#### Future Work:

All of the evidence for the involvement of these genes in the clinical features of the 22q13 deletion syndrome is circumstantial. Since some of the clinical features of the 22q13 deletion syndrome appear to have a parent of origin effect, it would be useful to study the expression pattern of these genes to detect possible imprinting effects.

#### Characterization of Unusual Patients:

##### Bisatellited Chromosome 22:

Three patients in this study were cytogenetically observed to have a bisatellited, dicentric chromosome 22 (figure 3-13). This type of chromosome is not formed by a simple rearrangement of chromosomal material, but a complex translocation involving both copies of chromosome 22. Deletions were identified in all of three of these patients indicating that there is a chromosomal breakage in 22q13. The p-arm of these bisatellited, dicentric chromosomes is duplicated and due to the repetitive content of all acrocentric p-arms this likely does not contribute to the phenotype of these patients. On





the other hand, duplication of the q-arm does result in clinical features. Three copies of 22q11 are sufficient to produce cat eye syndrome (Reiss, et al, 1985; Knoll, et al, 1995). Therefore it was of interest to determine if the phenotype in these patients was due to the deletion of 22q13, the duplication of 22q11 or a combination of both.

Of the three patients with this bisatellited, dicentric chromosome, patient 4 was extensively studied in collaboration with T. Babineau. Patient 4's bisatellited chromosome was found to be composed entirely of material from chromosome 22 with a duplication of the centromere and a deletion of the telomere. This patients showed some features that are observed with CES, including preauricular pits and developmental delay. However, all of the clinical features of this patient are also observed in the 22q13 deletion syndrome, including preauricular pits. Neither of the other patients with this unusual chromosomal rearrangement, 20 and 32, showed any features of CES. The deletion of 22q13 was confirmed in all three patients. However, no duplication of any of the genes in the CES critical region was detected. Since the most proximal PAC used from the CES critical region, which contains part of the pericentromeric region, is not duplicated the duplication in patient 4 is restricted to the centromeric and perhaps part of the pericentromeric region. The other two patients have not been verified for the duplication of the pericentromeric region and it might be of interest to determine the extent of the duplication in these patients. Thus the clinical features observed in these three patients with the bisatellited chromosome 22 are due to the deletion of 22q13 and not a duplication of 22q11.

It is interesting that 3/50 patients showed a more complex deletion involving a



bisatellited, dicentric chromosome. It would be beneficial to characterize the breakpoints in these patients in more detail to try to determine the mechanism by which this unusual chromosomal anomaly is formed. A recent abstract identifies a fourth patient with a similar structure of chromosome (Pulijall, et al, 2001). This patient has the key clinical features of the 22q13 deletion syndrome and thus may have a similar chromosome structure to the three patients in this thesis.

#### Patient with TAPVR and 22q13 deletion:

Another patient, 36, was suspected to have a duplication of the cat eye syndrome critical region due to the presence of TAPVR, an uncommon heart anomaly that is often associated with CES. The deletion of 22q13 region was cytogenetically detectable using G-banding. A medium-sized deletion was confirmed using microsatellite and FISH analysis. No duplication of the Cat Eye Syndrome critical region was detected in this patient through FISH analysis. This patient is being studied further in an attempt to detect a possible microduplication of the CES critical region. TAPVR has not been reported in any cases of the 22q13 deletion syndrome, and thus the genetic cause of the heart anomaly in this patient is unknown at this time.

#### Translocation Patient t(14;22):

Patient 2 was the product of the unbalanced segregation of a balanced familial translocation. The mother, grandmother, aunt and cousin of this patient all carry a balanced t(14;22) and are phenotypically normal, suggesting that there are no critical genes disrupted in these individuals. The patient however, inherited the derivative 22 and a normal chromosome 14 resulting in a deletion of the telomeric region of chromosome



22 and a duplication of the telomeric region of chromosome 14. The deletion present in this patient was too small to detect by microsatellite analysis. FISH using cosmid cN66C4, that contains the 5' end of the *SHANK3* gene, did not appear to be deleted in this patient. However FISH with cN85A3 confirmed that the 3' end of *SHANK3* was deleted. Since a published disruption of *SHANK3* causes clinical features observed in the 22q13 deletion syndrome (Bonaglia, et al, 2001), this implies that the breakpoint of this familial translocation must have occurred proximal to *SHANK3*, within cN66C4. This is supported by the fact that one cN66C4 signal was fainter than the other. This cosmid is approximately 45kb in size and about 35kb of this contains the exons of *SHANK3*. Therefore the breakpoint in this family is predicted to be within the proximal 10kb of this cosmid, leaving enough of the cosmid to allow for reduced hybridization. It is of importance to determine where the breakpoint in this familial translocation is. If the breakpoint does fall within *SHANK3*, then this contradicts the published translocation and raises a doubt about *SHANK3* as a candidate gene for the neurological abnormalities observed in the 22q13 deletion syndrome. However if the breakpoint is proximal to *SHANK*, then the location of the breakpoint would help to refine the 5' limit of the *SHANK3* gene including the regulatory region.

Both the patient 1/NT (Flint, et al, 1995) and the t(12;22) patient (Bonaglia, et al, 2001) have a mild phenotype compared to the rest of the 22q13 deletion syndrome patients. The developmental delay in both of these patients is relatively mild compared to other cases but neither patient has expressive speech. The t(12;22) patient does have hypotonia, whereas 1/NT does not. Patient 2 in this thesis has quite a severe phenotype:



she has profound developmental delay, delay of expressive speech, hypotonia and is wheelchair bound. The size of the deletion in this patient is not significantly larger than the deletion in patient 1/NT. Patient 2 is due to an unbalanced translocation that results in the deletion of chromosome 22q13 and the duplication of 14q. Whereas the t(12;22) patient is a balanced translocation and only two genes are interrupted, *SHANK3* and a gene of unknown function on chromosome 12. The difference in severity of phenotype between these two translocation patients may be due to the nature of the translocation, the fact that the milder patient has a balanced translocation and that the unbalanced patient is more severe.

#### JaA Mosaic Deletion:

Patient JaA was shown through cytogenetics to be mosaic for a deletion of 22q13. Most of the cells in JaA are of a normal karyotype (70%) and the remaining 30% of the cells have a deletion. Microsatellite analysis was not helpful in determining the size of the deletion in JaA since he has such a large proportion of normal cells. It is interesting that this patient shows such a severe phenotype and yet has such a high proportion of normal cells. JaA has severe developmental delay, delay of expressive speech, hypotonia, and several of the mild dysmorphic features associated with the 22q13 deletion syndrome. The severity in the phenotype may be due to a large proportion of cells in critical tissues show the deletion or at a critical time, whereas the blood has a higher proportion of the normal cells.

#### Interstitial Deletion (MRS):

Patient MRS was cytogenetically identified as having an interstitial deletion of





22q13. However a previous patient, 50, was also diagnosed cytogenetically as having an interstitial deletion but later was confirmed to have a terminal deletion of 22q13 based on the minisatellite D22S163. However MRS did show two copies of the minisatellite D22S163 confirming that this was not a terminal deletion. Several microsatellites also confirmed this. However this does not act as evidence against *SHANK3* being responsible for the neurological features observed in the 22q13 deletion syndrome. The distal breakpoint in this patient is found more proximal than the entire deletion in 19 patients which show the typical clinical features of the syndrome. Thus patient MRS is not a case of the “22q13 deletion syndrome”, but rather a different syndrome with similar clinical features. The proximal breakpoint of MRS is located in the same region as the breakpoint of largest 22q13 terminal deletions. Thus the MRS interstitial deletion is wholly contained within the largest terminal deletions 54 and 55. Some genes in this more proximal region of 22q13 must be responsible for the phenotype observed in MRS, and may contribute to the phenotype observed in the largest terminal deletions. MRS has developmental delay and absence of expressive speech (uses only two words at approximately 3 years). MRS is still very young, and her speech may improve as she ages. Part of the delay in her speech may be due to the severe developmental delay, or perhaps a problem with ear development. It would be interesting to observe her development to see if any more phenotypic differences develop. However, MRS does not have hypotonia or any other clinical features that are commonly observed in the 22q13 deletion syndrome. Abnormal neurological scans were noted in MRS but the findings do not correspond to those observed in any patients with the 22q13 deletion syndrome. MRS



also has difficulty walking due to problems with balance and not hypotonia. Although the clinical features of MRS are similar to those observed in the 22q13 deletion syndrome, MRS appears to have a different syndrome with a different genetic etiology.

Two other patients have been reported with interstitial deletions of 22q13. One of these patients was cytogenetically determined to have an interstitial deletion of 22q13.1 to 22q13.33, however no molecular studies were carried out (Romain, et al, 1990). The phenotype more closely matches that of the terminal 22q13 deletion syndrome than MRS with developmental delay, delay of expressive speech, hypotonia, overgrowth, ptosis, full cheeks, as well as several other minor dysmorphisms observed with the 22q13 deletion syndrome. Therefore Romain's patient may be a terminal deletion of 22q13 and not an interstitial deletion or may be another genetic syndrome that is related. The other patient (Fujita, et al, 2000) was confirmed to have an interstitial deletion by FISH analysis. However the cosmids used for FISH analysis on this patient have not been localized on the genomic sequence provided by the Sanger Center, and thus the precise location of this deletion has not been determined. Fujita's patient shows developmental delay, hypotonia and delay of expressive speech. However the absence of expressive speech in this patient is likely due to a bilateral profound hearing loss caused by an abnormal cochlea and vestibule malformations of the inner ear. This patient also has overgrowth, macrocephaly and several other mild dysmorphisms commonly observed with the 22q13 deletion syndrome. It would be important to further characterize the deletion in these two patients to determine if the deletion overlaps with the deletion observed in the patient MRS.

These few cases of interstitial deletions of 22q13 are interesting in that they have



some overlap of the clinical features, but do not quite seem to fit the phenotype of the 22q13 deletion syndrome. The interstitial deletions provide evidence that there is a second similar genetic syndrome that results from the deletion within the 22q13 region.

#### Future Work:

Patients with unusual chromosomal structure and rearrangements, are a goldmine of information. Any patient that cytogenetically shows an abnormality of chromosome 22q13 should be further characterized. Further characterization of the breakpoints in the three patients that show a bisatellited chromosome 22 may provide insight into their mechanism of origin. Comparing the cases with interstitial deletions to the terminal deletions of the 22q13 deletion syndrome may allow for phenotypic mapping of genes involved in the various features of both syndromes. Patients with the 22q13 deletion syndrome may also be deleted for the critical region for an interstitial deletion of 22q13, and thus may show phenotypic characteristics of both genetic syndromes. Further characterization of other interstitial deletions would help to refine the critical region for this syndrome, and provide a better clinical picture. Thus characterization of unusual patients in this study provided identification of an unusual bisatellited chromosome, and characterization of an interstitial deletion syndrome that is distinct from the 22q13 deletion syndrome.

#### Other Future Experiments:

The translocation t(12;22) (Bonaglia, et al, 2001) provides support to the hypothesis that *SHANK3* is responsible for the neurological abnormalities associated with the 22q13 deletion syndrome, but there is still the possibility that position effects and/or





the disruption of the gene on chromosome 12 may be responsible for the phenotype. However, in order to show definitively that *SHANK3* is responsible for the neurological abnormalities, a mutation in *SHANK3* would need to be found in a patient that shows a similar neurological phenotype but without a deletion of 22q13. This study is underway, examining a large sample of patients that show idiopathic developmental delay and delay or absence of expressive speech with little or no dysmorphic features. Many of the patients selected for this study are individuals that have tested negative for Fragile X, or Angelman syndrome. These patients are being screened for the presence of a small deletion that may remove a few exons of *SHANK3*, using multiple amplifiable probe hybridization (Sismani, et al, 2001). If no deletion is detected in these patients, then they will be screened for mutations in the coding region of *SHANK3* using denaturing high performance liquid chromatography (Oefner, and Underhill, 1995; Underhill, et al, 1997). Obtaining a mutation in *SHANK3* that causes disruption of the protein would provide the evidence needed to show that *SHANK3* is responsible for the neurological features observed in the 22q13 deletion syndrome.

The 22q13 deletion syndrome is not a widely recognized genetic syndrome for several reasons. The clinical features associated with this syndrome are not distinct. For example developmental delay and delay of expressive speech are found in many other disorders including Angelman's syndrome, FG syndrome, and Fragile X syndrome. In fact many of the cases in the literature, as well as in this study were originally tested for these other syndromes before the 22q13 deletion syndrome was tested for. The two





patients in Precht, et al (1998) were originally analyzed for DiGeorge/VCF syndrome and Angelman syndrome. Two of the three patients in Prasad, et al (2000) were originally thought to be DIG/VCFS based on their clinical features. However, FISH analysis for DGS/VCFS showed a deletion of the control probe in both patients. The patient published by DeVries, et al (2000), was originally characterized as a patient with FG syndrome. The dysmorphic features observed in the 22q13 deletion syndrome are mild and not present in all individuals and thus do not allow for a definitive diagnosis.

Second, chromosome 22 is a pale staining chromosome and thus it is difficult to identify small abnormalities. Many of the small 22q13 deletions are not detectable cytogenetically. Thus identification of the smaller deletions requires a specific molecular test such as FISH. Luckily, commercially available FISH probes to test for the DiGeorge/VCF syndrome deletion use a control probe in the 22q13 region. The Oncor probe, D22S39, is the most proximal control probe and will miss any deletions under 2Mb in size. Therefore this probe does not provide much assistance in the detection of the most terminal deletions. The control probe from Vysis, ARSA, is located approximately 150kb from the telomere and thus will detect almost all of the terminal deletions of 22q13. However this probe will miss the extremely small deletions. The other commercially available probe from Cytocell contains the most distal unique clones on chromosome 22, c202 and cN94H12, and therefore this probe will detect all terminal deletions. If a child presents with severe to profound developmental delay and delay or absence of expressive speech and there is no cytogenetic abnormality, then FISH with either the Vysis or Cytocell control probes may detect a cryptic deletion of 22q13.



Third, there is a lack of awareness regarding this syndrome in the medical community. DiGeorge/VCF syndrome are frequently referred to as “The 22 deletion syndrome” implying that it is the only disorder that results from a deletion within chromosome 22. Physicians and genetic counselors need to be made more aware of the clinical features of this syndrome, in order to better diagnose patients.



## Overall Conclusions:

1. *SHANK3* is presented as a candidate gene for the neurological abnormalities observed in the 22q13 deletion syndrome. This protein is a structural component of the synapse and might be sensitive to dosage. This gene is also expressed at high levels in the brain and shows a brain predominant transcript.
2. All of the patient deletions in this study that could be fully characterized are terminal deletions that result in the deletion of *SHANK3*. Considerable variation is observed in the sizes of the deletions and the larger deletions appear to have a higher incidence of several clinical features.
3. There are significantly more paternal deletions than maternal deletions. Several clinical features showed a higher incidence depending on the parent of origin of the deletion implying a possible imprinting effect.
4. The interstitial deletion, MRS, shares several clinical features with the 22q13 deletion syndrome but has a distinct, but overlapping syndrome compared to the 22q13 deletion syndrome. Several of the patients with larger deletions may show clinical features of both syndromes.
5. Several other genes in 22q13 may contribute to the various clinical features of the 22q13 deletion syndrome and need to be investigated further.







Figure 4-1: Protein components of the post-synaptic density of excitatory synapses. The network of proteins connecting the membrane receptors to the cytoskeleton is shown. The Shank proteins are linked in a head to head manner through an interaction of the ankyrin repeat domain with Sharpin. No interactions are known for the SH3 domain of Shank. The PDZ domain is responsible for the interaction of Shank with SAPAPs which connect to the metabotropic glutamate receptor. The Homer-binding domain binds to a homodimer of Homer which connects to the NMDA receptor. The SH3 binding domain binds to cortactin which joins to the cytoskeleton of the synapse. The SAM domain is responsible for the multimerization of the Shanks in a tail-to tail manner. Modified from Ehlers, 1999.

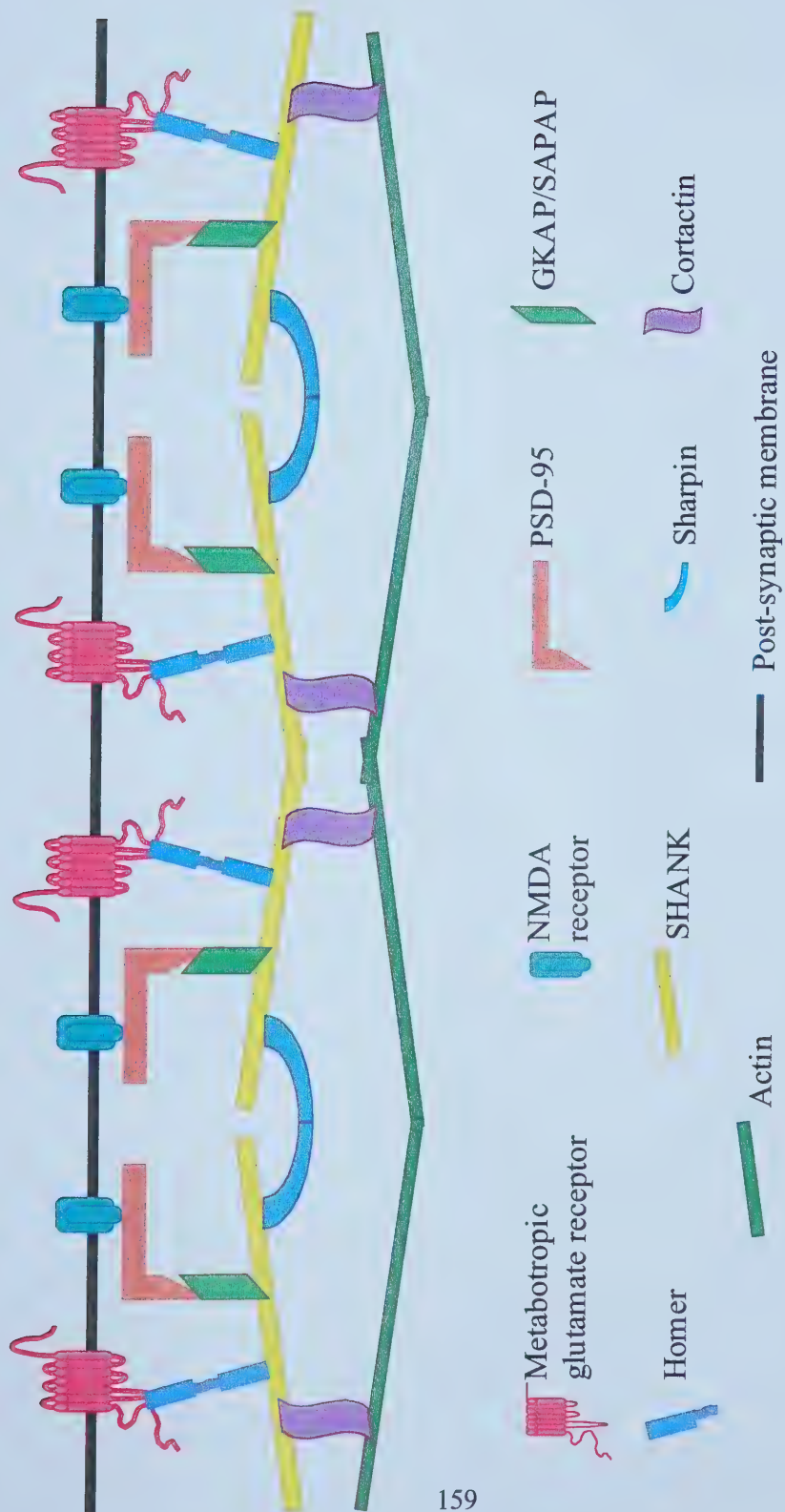






Figure 4-2: Genomic structure and predicted transcripts of *SHANK3*. A: Genomic structure of *SHANK3* showing each exon as a vertical bar. Dotted vertical bar is a predicted exon. Genomic sequence is represented by horizontal bar. Alternative 3' ends are identified by AAA. B: Predicted transcripts of *SHANK3*. First predicted transcript is 1623bp in size, believed to be 2kb transcript on Northern Blots. Second predicted transcript is 5784bp in size, believed to be 7.5kb brain predominant transcript on Northern Blots. Third predicted transcript is 6375bp in size, believed to be 8kb transcript on Northern Blots. Fourth predicted transcript is 7030bp in size, believed to be 10kb transcript on Northern Blots.

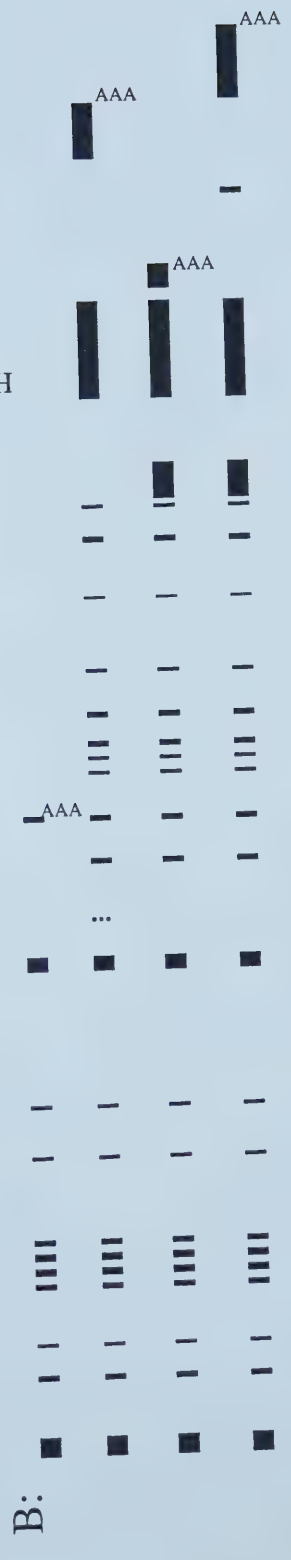
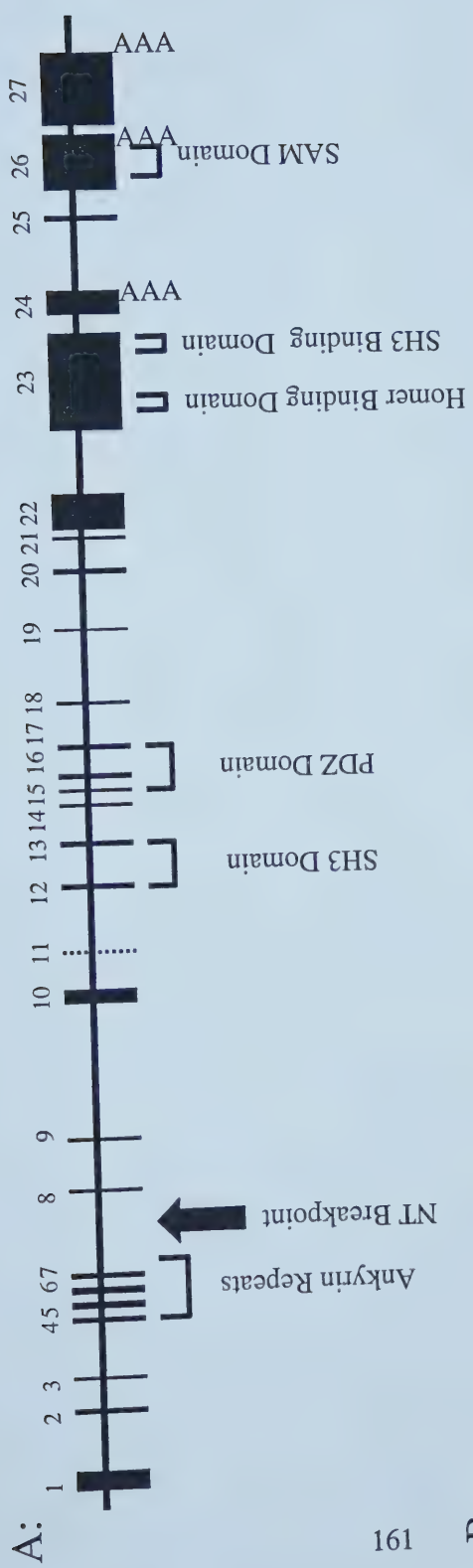
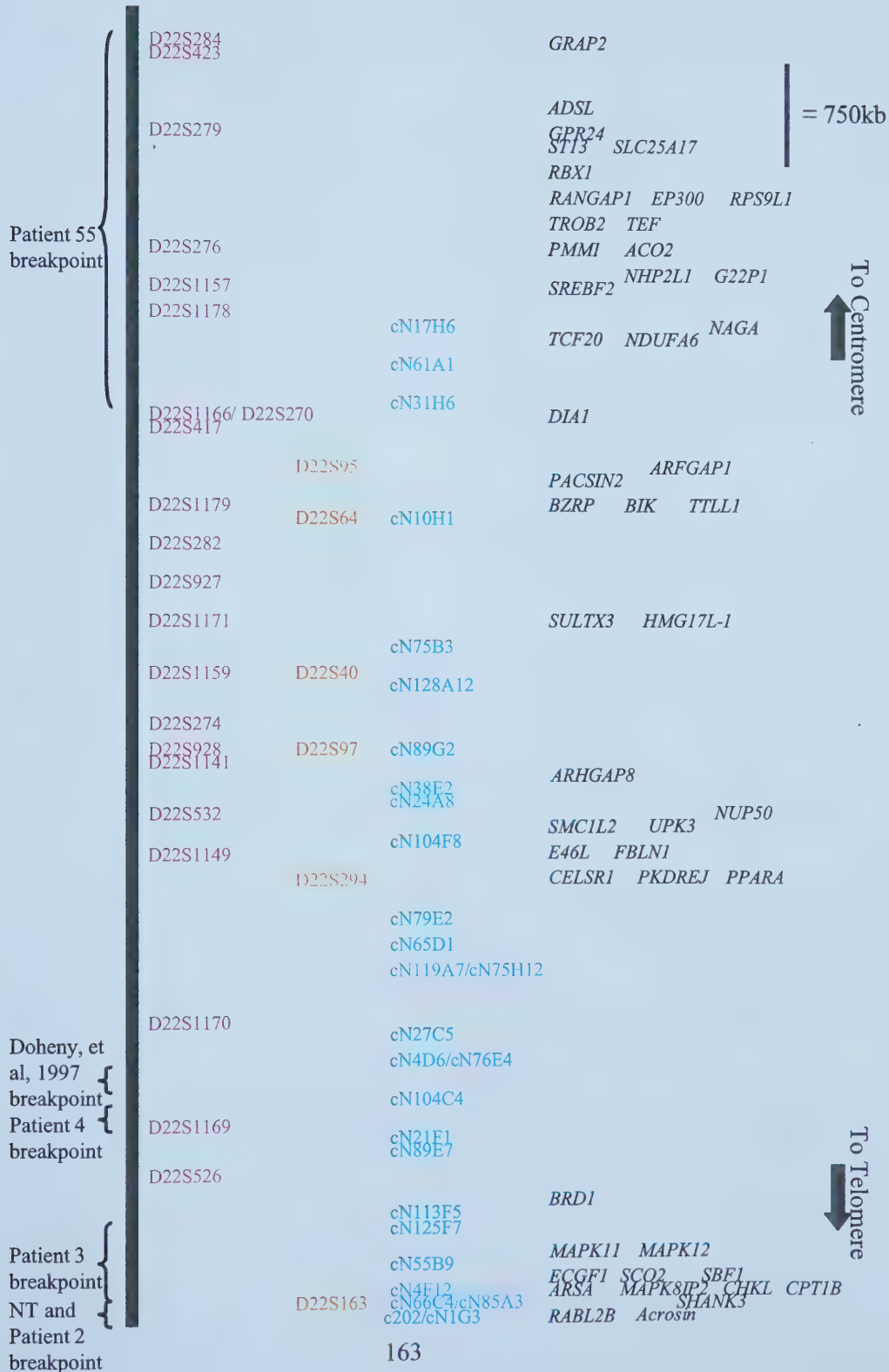








Figure 4-3: Diagram of region deleted in 22q13 deletion syndrome. Several key deletions are identified along the left side by showing region containing the breakpoint. Microsatellite and minisatellite markers are listed in purple and orange respectively. Cosmids used for FISH are listed in teal. Known genes in the region are listed in black on the right side. Positions of loci are roughly to scale. Note a cluster of genes at the terminal region of 22q13, this cluster also contains many predicted genes. This cluster of genes is followed by a gene poor region with no known genes and very few predicted genes.





## Bibliography

- Arakai DT, Sparkes RS (1963) Microtechniques for culturing leukocytes from whole blood. *Cytogenetics* 2:57-63.
- Armour JA, Jeffreys AJ (1991) STS for minisatellite MS607 (D22S163). *Nucleic Acids Res* 19:3158.
- Ballif BC, Kashork CD, Shaffer LG (2000) FISHing for mechanisms of cytogenetically defined terminal deletions using chromosome-specific subtelomeric probes. *Eur J Hum Genet* 8:764-70.
- Boeckers TM, Kreutz MR, Winter C, Zuschratter W, Smalla KH, Sanmarti-Vila L, Wex H, et al (1999a) Proline-rich synapse-associated protein-1/cortactin binding protein 1 (ProSAP1/CortBP1) is a PDZ-domain protein highly enriched in the postsynaptic density. *J Neurosci* 19:6506-18.
- Boeckers TM, Winter C, Smalla KH, Kreutz MR, Bockmann J, Seidenbecher C, Garner CC, et al (1999b) Proline-rich synapse-associated proteins ProSAP1 and ProSAP2 interact with synaptic proteins of the SAPAP/GKAP family. *Biochem Biophys Res Commun* 264:247-52.
- Bonaglia MC, Giorda R, Borgatti R, Felisari G, Gagliardi C, Selicorni A, Zuffardi O (2001) Disruption of the ProSAP2 gene in a t(12;22)(q24.1;q13.3) is associated with the 22q13.3 deletion syndrome. *Am J Hum Genet* 69:261-8.
- Carlson C, Sirotkin H, Pandita R, Goldberg R, McKie J, Wadey R, Patanjali SR, et al (1997) Molecular definition of 22q11 deletions in 151 velo-cardio-facial syndrome patients. *Am J Hum Genet* 61:620-9.
- Cassidy S, Cunay-Aygun M, Everman D, Winters J, Schwartz S (1999) Terminal 22q Deletion syndrome. In: XX David W. Smith Workshop on Malformations and Morphogenesis. 127.
- Craven SE, Brecht DS (1998) PDZ proteins organize synaptic signaling pathways. *Cell* 93:495-8.
- Cui X, De Vivo I, Slany R, Miyamoto A, Firestein R, Cleary ML (1998) Association of SET domain and myotubularin-related proteins modulates growth control. *Nat Genet* 18:331-7.



de Vries BB, Bitner-Glindzicz M, Knight SJ, Tyson J, MacDermont KD, Flint J, Malcolm S, et al (2000) A boy with a submicroscopic 22qter deletion, general overgrowth and features suggestive of FG syndrome. *Clin Genet* 58:483-7.

de Vries BBA, White SM, Knight SJL, Regan R, Homgray T, Young ID, Super M, McKeowan C, Splitt M, Quarrell OWJ, Trainer AH, Niermeijer MF, Malcolm S, Flint J, Hurst JA, Winter RM (2001) Clinical studies on submicroscopic subtelomeric rearrangements: a checklist. *J Med Genet* 38:145-50.

Doheny KF, McDermid HE, Harum K, Thomas GH, Raymond GV (1997) Cryptic terminal rearrangement of chromosome 22q13.32 detected by FISH in two unrelated patients. *J Med Genet* 34:640-4.

Dunham I, Shimizu N, Roe BA, Chisoe S, Hunt AR, Collins JE, Bruskiewich R, et al (1999) The DNA sequence of human chromosome 22. *Nature* 402:489-95.

Ehlers MD (1999) Synapse structure: glutamate receptors connected by the shanks. *Curr Biol* 9:R848-50.

Ewart AK, Morris CA, Atkinson D, Jin W, Sternes K, Spallone P, Stock AD, et al (1993) Hemizygoty at the elastin locus in a developmental disorder, Williams syndrome. *Nat Genet* 5:11-6.

Ewart AK, Jin W, Atkinson D, Morris CA, Keating MT (1994) Supravalvular aortic stenosis associated with a deletion disrupting the elastin gene. *J Clin Invest* 93:1071-7.

Ferguson BM, Fangman WL (1992) A position effect on the time of replication origin activation in yeast. *Cell* 68:333-9.

Fisher E, Scambler P (1994) Human haploinsufficiency--one for sorrow, two for joy. *Nat Genet* 7:5-7.

Flint J, Wilkie AO, Buckle VJ, Winter RM, Holland AJ, McDermid HE (1995) The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. *Nat Genet* 9:132-40.

Flint J, Bates GP, Clark K, Dorman A, Willingham D, Roe BA, Micklem G, et al (1997a) Sequence comparison of human and yeast telomeres identifies structurally distinct subtelomeric domains. *Hum Mol Genet* 6:1305-13.

Flint J, Thomas K, Micklem G, Raynham H, Clark K, Doggett NA, King A, et al (1997b) The relationship between chromosome structure and function at a human telomeric region. *Nat Genet* 15:252-7.



Footz TK, Brinkman-Mills P, Banting GS, Maier SA, Riaz MA, Bridgland L, Hu S, et al (2001) Analysis of the Cat Eye Syndrome Critical Region in Humans and the Region of Conserved Synteny in Mice: A Search for Candidate Genes at or near the Human Chromosome 22 Pericentromere. *Genome Res* 11:1053-70.

Frangiskakis JM, Ewart AK, Morris CA, Mervis CB, Bertrand J, Robinson BF, Klein BP, et al (1996) LIM-kinase1 hemizygosity implicated in impaired visuospatial constructive cognition. *Cell* 86:59-69.

Fujita Y, Mochizuki D, Mori Y, Nakamoto N, Kobayashi M, Omi K, Kodama H, et al (2000) Girl with accelerated growth, hearing loss, inner ear anomalies, delayed myelination of the brain, and del(22)(q13.1q13.2). *Am J Med Genet* 92:195-9.

Ghaffari SR, Boyd E, Tolmie JL, Crow YJ, Trainer AH, Connor JM (1998) A new strategy for cryptic telomeric translocation screening in patients with idiopathic mental retardation. *J Med Genet* 35:225-33.

Goizet C, Excoffier E, Taine L, Taupiac E, El Moneim AA, Arveiler B, Bouvard M, et al (2000) Case with autistic syndrome and chromosome 22q13.3 deletion detected by FISH. *Am J Med Genet* 96:839-44.

Greenberg F, Elder FF, Haffner P, Northrup H, Ledbetter DH (1988) Cytogenetic findings in a prospective series of patients with DiGeorge anomaly. *Am J Hum Genet* 43:605-11.

Gustincich S, Manfioletti G, Del Sal G, Schneider C, Carninci P (1991) A fast method for high-quality genomic DNA extraction from whole human blood. *Biotechniques* 11:298-300, 302.

Hadjantonakis AK, Sheward WJ, Harmar AJ, de Galan L, Hoovers JM, Little PF (1997) Celsr1, a neural-specific gene encoding an unusual seven-pass transmembrane receptor, maps to mouse chromosome 15 and human chromosome 22qter. *Genomics* 45:97-104.

Hadjantonakis AK, Formstone CJ, Little PF (1998) mCelsr1 is an evolutionarily conserved seven-pass transmembrane receptor and is expressed during mouse embryonic development. *Mech Dev* 78:91-5.

Herman GE, Greenberg F, Ledbetter DH (1988) Multiple congenital anomaly/mental retardation (MCA/MR) syndrome with Goldenhar complex due to a terminal del(22q). *Am J Med Genet* 29:909-15.

Hughes J, Ward CJ, Aspinwall R, Butler R, Harris PC (1999) Identification of a human homologue of the sea urchin receptor for egg jelly: a polycystic kidney disease-like protein. *Hum Mol Genet* 8:543-9.





Hunter AG, Ray M, Wang HS, Thompson DR (1977) Phenotypic correlations in patients with ring chromosome 22. *Clin Genet* 12:239-49.

Jee C, Na Y, Ahnn J (2001) Characterization of the *shn-1*, shank homologue in *C. elegans*. 13<sup>th</sup> International *C. elegans* meeting: 418.

Kirshenbaum G, Chmura M, Rhone DP (1988) Long arm deletion of chromosome 22. *J Med Genet* 25:780.

Knight SJ, Horsley SW, Regan R, Lawrie NM, Maher EJ, Cardy DL, Flint J, et al (1997) Development and clinical application of an innovative fluorescence in situ hybridization technique which detects submicroscopic rearrangements involving telomeres. *Eur J Hum Genet* 5:1-8.

Knight SJ, Regan R, Nicod A, Horsley SW, Kearney L, Homfray T, Winter RM, et al (1999) Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet* 354:1676-81.

Knight SJ, Flint J (2000a) Perfect endings: a review of subtelomeric probes and their use in clinical diagnosis. *J Med Genet* 37:401-9.

Knight SJ, Flint J (2000b) Screening chromosome ends for learning disability. *Bmj* 321:1240.

Knight SJ, Lese CM, Precht KS, Kuc J, Ning Y, Lucas S, Regan R, et al (2000c) An optimized set of human telomere clones for studying telomere integrity and architecture. *Am J Hum Genet* 67:320-32.

Knoll JH, Asamoah A, Pletcher BA, Wagstaff J (1995) Interstitial duplication of proximal 22q: phenotypic overlap with cat eye syndrome. *Am J Med Genet* 55:221-4.

Kreysing J, von Figura K, Gieselmann V (1990) Structure of the arylsulfatase A gene. *Eur J Biochem* 191:627-31.

Lechner C, Zahalka MA, Giot JF, Moller NP, Ullrich A (1996) ERK6, a mitogen-activated protein kinase involved in C2C12 myoblast differentiation. *Proc Natl Acad Sci U S A* 93:4355-9.

Ledbetter DH (1992) Minireview: cryptic translocations and telomere integrity. *Am J Hum Genet* 51:451-6.



- Leegwater PA, Yuan BQ, van der Steen J, Mulders J, Konst AA, Boor PK, Mejaski-Bosnjak V, et al (2001) Mutations of MLC1 (KIAA0027), encoding a putative membrane protein, cause megalencephalic leukoencephalopathy with subcortical cysts. *Am J Hum Genet* 68:831-8.
- Li TC, Hu DY, Tang CS (1996) [Mitogen-activated protein kinase involved in intracellular signal transduction]. *Sheng Li Ke Xue Jin Zhan* 27:258-60.
- Lim S, Naisbitt S, Yoon J, Hwang JI, Suh PG, Sheng M, Kim E (1999) Characterization of the Shank family of synaptic proteins. Multiple genes, alternative splicing, and differential expression in brain and development. *J Biol Chem* 274:29510-8.
- Lim S, Sala C, Yoon J, Park S, Kuroda S, Sheng M, Kim E (2001) Sharpin, a novel postsynaptic density protein that directly interacts with the shank family of proteins. *Mol Cell Neurosci* 17:385-97.
- MacDonald HR, Wevrick R (1997) The necdin gene is deleted in Prader-Willi syndrome and is imprinted in human and mouse. *Hum Mol Genet* 6:1873-8.
- MacLean JE, Teshima IE, Szatmari P, Nowaczyk MJ (2000) Ring chromosome 22 and autism: report and review. *Am J Med Genet* 90:382-5.
- McDermid HE, Duncan AM, Brasch KR, Holden JJ, Magenis E, Sheehy R, Burn J, et al (1986) Characterization of the supernumerary chromosome in cat eye syndrome. *Science* 232:646-8.
- McQuillin A, Kalsi G, Moorey H, Lamb G, Mayet A, Questead D, Gurling H (2001) Exon 11 of the chromosome 22 gene *WKL1*, previously shown to exhibit a base pair mutation in rare cases of schizophrenia, was found to contain an insertion/deletion polymorphism, affecting the receptor channel amino acid sequence. *Am J Hum Genet* 69:A2153.
- McTaggart KE, Budarf LM, Driscoll DA, Emanuel BS, Ferreira P, McDermid HE (1998) Cat eye syndrome chromosome breakpoint clustering: identification of two intervals also associated with 22q11 deletion syndrome breakpoints. *Cytogenet Cell Genet* 81:222-8.
- Meyer J, Huberth A, Ortega G, Syagailo YV, Jatzke S, Mossner R, Strom TM, et al (2001) A missense mutation in a novel gene encoding a putative cation channel is associated with catatonic schizophrenia in a large pedigree. *Mol Psychiatry* 6:302-6.
- Monaco AP (1996) Human genetics: dissecting Williams syndrome. *Curr Biol* 6:1396-8.



Morrow B, Goldberg R, Carlson C, Das Gupta R, Sirotkin H, Collins J, Dunham I, et al (1995) Molecular definition of the 22q11 deletions in velo-cardio-facial syndrome. *Am J Hum Genet* 56:1391-403.

Naisbitt S, Kim E, Tu JC, Xiao B, Sala C, Valtschanoff J, Weinberg RJ, et al (1999) Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin. *Neuron* 23:569-82.

Narahara K, Takahashi Y, Murakami M, Tsuji K, Yokoyama Y, Murakami R, Ninomiya S, et al (1992) Terminal 22q deletion associated with a partial deficiency of arylsulphatase A. *J Med Genet* 29:432-3.

Naritomi K, Hirayama K (1988) Determination of the breakpoints and the parental origin of a ring 22 chromosome: an analysis by high-resolution banding technique, quinacrine and silver stainings. *Jinrui Idengaku Zasshi* 33:67-73.

National Institutes of Health and Institute of Molecular Medicine collaboration (1996) A complete set of human telomeric probes and their clinical application. *Nat Genet* 14:86-9.

Nesslinger NJ, Gorski JL, Kurczynski TW, Shapira SK, Siegel-Bartelt J, Dumanski JP, Cullen RF, Jr., et al (1994) Clinical, cytogenetic, and molecular characterization of seven patients with deletions of chromosome 22q13.3. *Am J Hum Genet* 54:464-72.

Nishino I, Spinazzola A, Hirano M (2001) MNGIE: from nuclear DNA to mitochondrial DNA. *Neuromuscul Disord* 11:7-10.

Nomura N, Miyajima N, Sazuka T, Tanaka A, Kawarabayasi Y, Sato S, Nagase T, et al (1994) Prediction of the coding sequences of unidentified human genes. I. The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1 (supplement). *DNA Res* 1:47-56.

Oefner PJ, Underhill PA (1995) Comparative DNA sequencing by denaturing high-performance liquid chromatography (DHPLC). *Am J Hum Genet* 57:A266.

Ofir R, Wong AC, McDermid HE, Skorecki KL, Selig S (1999) Position effect of human telomeric repeats on replication timing. *Proc Natl Acad Sci U S A* 96:11434-9.

Online Mendelian Inheritance in Man, OMIM (TM). Center for Medical Genetics, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine (Bethesda, MD), 1999. World Wide Web URL: [www.ncbi.nlm.gov/omim/](http://www.ncbi.nlm.gov/omim/)





Papadopoulou LC, Sue CM, Davidson MM, Tanji K, Nishino I, Sadlock JE, Krishna S, et al (1999) Fatal infantile cardioencephalomyopathy with COX deficiency and mutations in SCO2, a COX assembly gene. *Nat Genet* 23:333-7.

Phelan MC, Rogers RC, Stevenson RE (1988) A de novo terminal deletion of 22q. *Am J Hum Genet* 43:A118.

Phelan MC, Thomas GR, Saul RA, Rogers RC, Taylor HA, Wenger DA, McDermid HE (1992) Cytogenetic, biochemical, and molecular analyses of a 22q13 deletion. *Am J Med Genet* 43:872-6.

Phelan MC, Rogers RC, Saul RA, Stapleton GA, Sweet K, McDermid H, Shaw SR, et al (2001) 22q13 deletion syndrome. *Am J Med Genet* 101:91-9.

Praphanphoj V, Goodman BK, Thomas GH, Raymond GV (2000) Cryptic subtelomeric translocations in the 22q13 deletion syndrome. *J Med Genet* 37:58-61.

Prasad C, Prasad AN, Chodirker BN, Lee C, Dawson AK, Jocelyn LJ, Chudley AE (2000) Genetic evaluation of pervasive developmental disorders: the terminal 22q13 deletion syndrome may represent a recognizable phenotype. *Clin Genet* 57:103-9.

Precht KS, Lese CM, Spiro RP, Huttenlocher PR, Johnston KM, Baker JC, Christian SL, et al (1998) Two 22q telomere deletions serendipitously detected by FISH. *J Med Genet* 35:939-42.

Pulijall VR, Gogineni SK, Gross SJ, Iacoboni D, Marion RW, Shanske A (2001) Molecular characterization of a satellited chromosome 22 in a patient with mental retardation and no dysmorphic features. *Am J Hum Genet* 69:A844.

Redecker P, Gundelfinger ED, Boeckers TM (2001) The cortactin-binding postsynaptic density protein proSAP1 in non- neuronal cells. *J Histochem Cytochem* 49:639-48.

Reiss JA, Weleber RG, Brown MG, Bangs CD, Lovrien EW, Magenis RE (1985) Tandem duplication of proximal 22q: a cause of cat-eye syndrome. *Am J Med Genet* 20:165-71.

Riegel M, Baumer A, Wisser J, Acherman J, Schinzel A (2000) Prenatal diagnosis of mosaicism for a del(22)(q13). *Prenat Diagn* 20:76-9.

Romain DR, Goldsmith J, Cairney H, Columbano-Green LM, Smythe RH, Parfitt RG (1990) Partial monosomy for chromosome 22 in a patient with del(22)(pter---q13.1::q13.33---qter). *J Med Genet* 27:588-9.





- Saccone S, De Sario A, Della Valle G, Bernardi G (1992) The highest gene concentrations in the human genome are in telomeric bands of metaphase chromosomes. *Proc Natl Acad Sci U S A* 89:4913-7.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* 2<sup>nd</sup> Edition. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.
- Sambrook J, Russell DW (2001) *Molecular Cloning: A Laboratory Manual* 3<sup>rd</sup> Edition. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.
- Schinzel A, Schmid W, Fraccaro M, Tiepolo L, Zuffardi O, Opitz JM, Lindsten J, et al (1981) The "cat eye syndrome": dicentric small marker chromosome probably derived from a no.22 (tetrasomy 22pter to q11) associated with a characteristic phenotype. Report of 11 patients and delineation of the clinical picture. *Hum Genet* 57:148-58.
- Schinzel AA, Basaran S, Bernasconi F, Karaman B, Yuksel-Apak M, Robinson WP (1994) Maternal uniparental disomy 22 has no impact on the phenotype. *Am J Hum Genet* 54:21-4.
- Schroder K, Schuffenhauer S, Seidel H, Bartsch O, Blin N, Hinkel GK, Schmitt H (1998) Deletion mapping by FISH with BACs in patients with partial monosomy 22q13. *Hum Genet* 102:557-61.
- Shaikh TH, Kurahashi H, Saitta SC, O'Hare AM, Hu P, Roe BA, Driscoll DA, et al (2000) Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Hum Mol Genet* 9:489-501.
- Shinoura N, Shamraj OI, Hugenholz H, Zhu JG, McBlack P, Warnick R, Tew JJ, et al (1995) Identification and partial sequence of a cDNA that is differentially expressed in human brain tumors. *Cancer Lett* 89:215-21.
- Shprintzen RJ, Goldberg RB, Young D, Wolford L (1981) The velo-cardio-facial syndrome: a clinical and genetic analysis. *Pediatrics* 67:167-72.
- Sismani C, Armour JA, Flint J, Girgalli C, Regan R, Patsalis PC (2001) Screening for subtelomeric chromosome abnormalities in children with idiopathic mental retardation using multiprobe telomeric FISH and the new MAPH telomeric assay. *Eur J Hum Genet* 9:527-32.
- Slavotinek A, Maher E, Gregory P, Rowlandson P, Huson SM (1997) The phenotypic effects of chromosome rearrangement involving bands 7q21.3 and 22q13.3. *J Med Genet* 34:857-61.



Slavotinek A, Rosenberg M, Knight S, Gaunt L, Fergusson W, Killoran C, Clayton-Smith J, et al (1999) Screening for submicroscopic chromosome rearrangements in children with idiopathic mental retardation using microsatellite markers for the chromosome telomeres. *J Med Genet* 36:405-11.

Smith DP, Floyd M, Say B (1996) Detection of a familial cryptic translocation by fluorescent in situ hybridisation. *J Med Genet* 33:84.

Smith A, Jauch A, St Heaps L, Robson L, Kearney B (2000) Unbalanced translocation t(15;22) in "severe" Prader-Willi syndrome. *Ann Genet* 43:125-30.

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98:503-17.

Tachdjian G, Muti C, Gaudelus J, Druart L, Martin B, Tamboise E, Nessmann C (1992) Unbalanced karyotype due to adjacent 1 segregation of t(11;22)(q23.3;q13.2). *Ann Genet* 35:231-3.

Tassabehji M, Metcalfe K, Fergusson WD, Carette MJ, Dore JK, Donnai D, Read AP, et al (1996) LIM-kinase deleted in Williams syndrome. *Nat Genet* 13:272-3.

Ton CC, Hirvonen H, Miwa H, Weil MM, Monaghan P, Jordan T, van Heyningen V, et al (1991) Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. *Cell* 67:1059-74.

Tu JC, Xiao B, Naisbitt S, Yuan JP, Petralia RS, Brakeman P, Doan A, et al (1999) Coupling of mGluR/Homer and PSD-95 complexes by the Shank family of postsynaptic density proteins. *Neuron* 23:583-92.

Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, et al (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* 7:996-1005.

van der Knaap MS, Barth PG, Stroink H, van Nieuwenhuizen O, Arts WF, Hoogenraad F, Valk J (1995) Leukoencephalopathy with swelling and a discrepantly mild clinical course in eight children. *Ann Neurol* 37:324-34.

Wang PJ, McCarrey JR, Yang F, Page DC (2001) An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* 27:422-6.

Watt JL, Olson IA, Johnston AW, Ross HS, Couzin DA, Stephen GS (1985) A familial pericentric inversion of chromosome 22 with a recombinant subject illustrating a 'pure' partial monosomy syndrome. *J Med Genet* 22:283-7.



Wong AC, Ning Y, Flint J, Clark K, Dumanski JP, Ledbetter DH, McDermid HE (1997) Molecular characterization of a 130-kb terminal microdeletion at 22q in a child with mild mental retardation. *Am J Hum Genet* 60:113-20.

Wong AC (1998) Transcriptional mapping in a terminal microdeletion of human chromosome 22q. PhD Thesis Department of Biological Sciences. University of Alberta.

Wong AC, Shkolny D, Dorman A, Willingham D, Roe BA, McDermid HE (1999) Two novel human RAB genes with near identical sequence each map to a telomere-associated region: the subtelomeric region of 22q13.3 and the ancestral telomere band 2q13. *Genomics* 59:326-34.

Yamazaki N, Shinohara Y, Shima A, Yamanaka Y, Terada H (1996) Isolation and characterization of cDNA and genomic clones encoding human muscle type carnitine palmitoyltransferase I. *Biochim Biophys Acta* 1307:157-61.

Yong YP, Knight LA, Yong MH, Lam S, Ho LY (1997) Partial monosomy for chromosome 22 in a girl with mental retardation. *Singapore Med J* 38:85-6.

Zackowski JL, Raffel LJ, McDaniel LD, Schwartz S (1990) A paternal balanced translocation [t(7;22)(q32;q13.3)] leading to reciprocal unbalanced karyotypes in two consecutive pregnancies. *Ann Genet* 33:113-6.

Zwaigenbaum L, Siegel-Bartelt J, Teshima I, Ho C (1990) Two patients with 22q13.3 deletions have similar facies and developmental patterns. *Am J Hum Genet* 47:A45.



## Appendix A: % GC Formula for Calculating $T_m$ of Primers

The melting temperature of primers was calculated using the following formula:

$$T_m = 67.5 + 34 (\%GC) - (395/\text{number of bases})$$

Where: %GC = number of bases that are G or C/ total number of bases in primer  
number of bases = total number of bases in primer





**Appendix B:** Questions used in analysis of phenotype of patients with the 22q13 deletion syndrome.

1. What are the IQ scores when broken down into different components?
2. What are the number of words spoken by patient?
3. What is the comprehension of spoken words?
4. Which patients use gestures to communicate? Which ones use few gestures?
5. Which patients have delayed motor milestones? Which ones were do not?
6. Which patients have hypotonia, which ones do not? Is the hypotonia global?
7. Which patients are able to walk unaided? Which are able to walk using a walker?  
Which are able to walk with parental assistance? Which ones are unable to walk?
8. Which patients have excessive growth? Which patients have normal growth?  
Which patients have undergrowth?
9. Which patients have organ malformations? What organs are affected?
10. Which patients have had an MRI? Which patients had an abnormal MRI?
11. What facial dysmorphisms does the patient have?
12. Which patients have abnormal sleep patterns/sleep disruption?
13. Which patients have feeding problems?
14. Which patients have good eye contact?
15. Which patients interact well with others?
16. Which patients have a normal attention span?
17. Which patients have a high pain tolerance?
18. Which patients use non-verbal vocalization to communicate?



19. Which patients have hearing defects?
20. Which patients are able to use communication aid devices?
21. Which patients have an abnormal gait?
22. Which patients have a larger than normal head circumference? Which ones have normal head circumferences? Which ones have smaller than normal head circumferences?
23. Which patients have abnormal hands or feet? What are these abnormalities?
24. Which patients have vesicoureteral reflux?
25. Which patients are prone to urogenital infections?
26. Which patients are prone to ear infections?
27. Which patients suffer from seizures?
28. Which patients have intestinal problems?
29. Which patients had pregnancy/neonatal complications?
30. Which patients have balance problems?
31. Which patients have kidney malformations?



**Appendix C:** Example of Linear Regression Analysis Output.

Data was coded with numbers to allow for analysis in a Microsoft Excel Spreadsheet. Linear regression was carried out one clinical feature (dependent variable) at a time in comparison with the size of the deletion (independent variable).

The important numbers for the analysis of this table are bolded. The Multiple R (0.6089) is equal to the Coefficient of Correlation. If the Coefficient of Correlation is +1 then there is a perfect positive correlation. If the Coefficient of Correlation is -1 then there is a perfect negative correlation. If the Coefficient of Correlation is 0 then no correlation exists. The closer the Coefficient of Correlation is to +1 or -1, the stronger the correlation between the two variables. The Significance F (0.0044) determines whether or not the Coefficient of Correlation is statistically significant or not. In this analysis if the Significance F is below 0.05 then the result is statistically significant.

Regression Statistics	
Multiple R	<b>0.6089</b>
R Square	0.3708
Adjusted R Square	0.3358
Standard Error	4.5358
Observations	20

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	218.2181	218.2181	10.6065	<b>0.0044</b>
Residual	18	370.331	20.5739		
Total	19	588.55			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	32.9803	5.8097	5.6767	2.199E-05	20.7745	45.1861	20.774	45.186
Size of Deletion (Mb)	-2.515	0.7725	-3.2567	0.0043	-4.1389	-0.8929	-4.138	-0.8929



**Appendix D:** Example of Contingency Table and Chi Square Analysis Output

Tables were created with the data as follows:

	Clinodactyly	Normal	Total
PAT deletion	1	19	20
MAT deletion	3	7	10
Total	4	26	30

Contingency tables were created to obtain the expected values for Chi Square analysis. The contingency table makes the assumption that the null hypothesis is that the two variables are independent from each other. In this case the null hypothesis would be that the parent of origin of the deletion has no affect on the clinical feature of clinodactyly.

	Clinodactyly	Normal	Total
PAT deletion	$(20*4)/30$	$(20*26)/30$	20
MAT deletion	$(10*4)/30$	$(10*26)/30$	10
Total	4	26	30

The numbers obtained for the above sample contingency table would be as follows:

	Clinodactyly	Normal	Total
PAT deletion	2.67	17.33	20
MAT deletion	1.33	6.67	10
Total	4	26	30





Chi square analysis was carried out using the numbers in the contingency table as the expected numbers. The basic formula for Chi Square is the sum of all  $(\text{observed}-\text{expected})^2/\text{expected}$ , and in this example would be **3.606**. The number of degrees of freedom for a contingency table is  $(\text{number of rows}-1)(\text{number of columns}-1)$ , and in this example would be one degree of freedom. A p-value can be obtained from a Critical Chi Square Value table, and in this example would be **0.056**.



**Appendix E:** Web Sites and Computer Software Used. List of Web Sites and URLs for web sites frequently accessed for this thesis. List of Software frequently used for this thesis.

BLAST (Basic Local Alignment Search Tool)

<http://www.ncbi.nlm.nih.gov/BLAST>

Deletion 22q13 Website

<http://www.22q13.com>

DNA Strider Computer Analysis Software

Entrez PubMed

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>

Genome Database

<http://gdbwww.gdb.org>

Genome Database Simple Search

<http://www.webace.sanger.ac.uk>

GENSCAN

<http://genes.mit.edu/GENSCAN.html>

Genetool Computer Software from Biotools Inc.

Human DNA Genomic Sequencing

[http://www.genome.ou.edu/hum\\_totals.html](http://www.genome.ou.edu/hum_totals.html)

Pip Maker

<http://bio.cse.psu.edu/pipmaker>

Primer3

[http://www.genome.wi.mit.edu/cgi-bin/primer/primer3\\_www.cgi](http://www.genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi)

RepeatMasker

<http://ftp.genome.washington.edu/cgi-bin/RepeatMasker>

Research Genetics

<http://www.resgen.com>



Sanger Centre

<http://www.sanger.ac.uk>

UniGene

<http://www.ncbi.nlm.nih.gov/UniGene>

















University of Alberta Library



0 1620 1520 5220

**B45570**